ON DIVISION ALGEBRAS*

BY

J. H. M. WEDDERBURN

§ 1. The object of this paper is to develop some of the simpler properties of division algebras, that is to say, linear associative algebras in which division is possible by any element except zero.

The determination of all such algebras in a given field is one of the most interesting problems in the theory of linear algebras. Early in the development of the subject, Frobenius showed that quaternions and its subalgebras form the only division algebras in the field of real numbers and, with the exception of the single theorem that there is no non-commutative division algebra in a finite field, no further definite result of importance was known till Dickson discovered the algebra referred to in § 4.

It is shown in the present paper that the Dickson algebra is the only non-commutative algebra of order 9 so that the only division algebras of order not greater than 9 are (i) the Dickson algebras of order 4 and 9, (ii) the ordinary commutative fields, (iii) algebras of order 8 which reduce to a Dickson algebra of order 4 when the field is extended to include those elements of the algebra which are commutative with every other element.

§ 2. Lemma 1. *If B is a subalgebra of order b in a division algebra A of order a, there exists a complex C of order c such that*

$$A = BC, \qquad a = bc.$$

Denoting elements of B by y with appropriate suffixes, let x_2 be an element of A which does not lie in B; the order of the complex $B + Bx_2$ is then $2b$ as otherwise there would be a relation of the form $y_1 + y_2 x_2 = 0$, $(y_2 \neq 0)$, which would lead to $x_2 = -y_2^{-1} y_1 < B$. Similarly, if $x_3 \nless B + Bx_2$, the order of $B + Bx_2 + Bx_3$ is $3b$ since otherwise there would be a relation of the form $y_1 + y_2 x_2 + y_3 x_3 = 0$, $(y_3 \neq 0)$, which would lead to

$$x_3 = -y_3^{-1} y_1 - y_3^{-1} y_2 x_2 < B + Bx_2.$$

Since the basis of A is finite, the truth of the lemma follows by an easy induction.

* Presented to the Society, February 28, 1920.

LEMMA 2. *If a polynomial*[*] $a_0 \xi^n + a_1 \xi^{n-1} + \cdots + a_n$ *in a scalar variable* ξ *is divided on the right and left by* $\xi - b$, *the remainders are* $a_0 b^n + a_1 b^{n-1} + \cdots + a_n$ *and* $b^n a_0 + b^{n-1} a_1 + \cdots + a_n$ *respectively.*

The proof of this lemma is exactly the same as in ordinary algebra, due care being taken to distinguish between multiplication on the right and on the left.

A factor which divides a polynomial on the right (left) will be referred to as a R.F. (L.F.).

LEMMA 3. *If*

$$A = a_0 \xi^m + a_1 \xi^{m-1} + \cdots + a_m$$

and

$$B = b_0 \xi^n + b_1 \xi^{n-1} + \cdots + b_n$$

are polynomials in a scalar variable ξ, *there exists a highest common right-hand factor* (*H.C.R.F.*) C_1 *and a highest common left-hand factor* (*H.C.L.F.*) C_2 *and polynomials* L_1, M_1, L_2, M_2 *such that*

$$L_1 A + M_1 B \equiv C_1, \qquad AL_2 + BM_2 \equiv C_2.$$

If $n \leqq m$, we can, by right-hand division, determine polynomials Q_1 and R_1 in ξ such that $A \equiv Q_1 B + R_1$, where R_1 is of lower degree in ξ than B. Obviously any C.R.F. of B and R_1 is a R.F. of A; we can therefore proceed with the proof exactly as in ordinary algebra.

The theory of linear factors of a polynomial in a scalar variable is by no means so simple as in commutative algebras. Their properties depend mainly on the following considerations. Let $A = BC$ be a polynomial in ξ expressed as the product of two polynomial factors B and C, and suppose that $\xi - x$ is a right factor of A but not of C; we have then $C = Q_1(\xi - x) + R$, where Q_1 is a polynomial and R is independent of ξ and is not zero. Multiplying by B we get $A = BQ_1(\xi - x) + BR$, whence $\xi - x$ is a R.F. of BR so that we can set $BR = Q_2(\xi - x)$ or $B = Q_2 R^{-1}(\xi - RxR^{-1})$, i.e., $\xi - RxR^{-1}$ is a R.F. of B.

A case of some importance arises when the algebra is quadrate i.e., where scalars are the only elements commutative with every element of the algebra, and when $A = 0$ is the reduced equation of this algebra. Regarding these algebras we have the following

LEMMA 4. *If* $\phi(x_1) = 0$ *is the reduced equation of an element of a quadrate division algebra, then, if* p *is the degree of* ϕ, *the scalar polynomial* $\phi(\xi)$ *can be expressed rationally as the product of* p *linear factors which may be permuted cyclically.*

Since $\phi(x_1) = 0$, $\xi - x_1$ is both a right and a left factor of $\phi(\xi)$, and so

[*] Throughout this paper all elements such as a_0, a_1, $\cdots$ are to be considered as belonging to a division algebra unless the contrary is stated explicitly.

is also $\xi - x'$ if x' is any transform of x_1. Let $\phi(\xi) = B(\xi - x_1)$, then, if x' is a transform which is not equal to x_1, we have

$$\phi(\xi) = B(\xi - x_1) = B(\xi - x') + B(x' - x_1);$$

hence, as above, $\xi - x_2 \equiv \xi - (x' - x_1)x'(x' - x_1)^{-1}$ is a R.F. of B. Similarly if $B = B'(\xi - x_2)$, and x'' is a transform of x_1 such that $\xi - x''$ is not a R.F. of $(\xi - x_2)(\xi - x_1)$, we find as above that

$$\xi - x_3 \equiv \xi - Rx''R^{-1}, \qquad R = x''^2 - (x_2 + x_1)x'' + x_2 x_1 \neq 0,$$

is a R.F. of B'; and so on. Continuing this process we get finally

$$\phi(\xi) = C(\xi - x_m)(\xi - x_{m-1}) \cdots (\xi - x_2)(\xi - x_1) \equiv CD \quad (m \leqq p),$$

where, if y is any transform of x_1, then $\xi - y$ is a R.F. of D. If therefore $D \equiv \xi^m + \alpha_1 \xi^{m-1} + \cdots + \alpha_m$, then

$$(1) \qquad y^m + \alpha_1 y^{m-1} + \cdots + \alpha_m = 0$$

for every transform y of x_1. If the α's are not all scalar, let z be an element which is not commutative with at least one of them and let $\alpha_i' = z\alpha_i z^{-1}$. Since (1) is satisfied by every transform of x_1, it follows that every transform also satisfies

$$(2) \qquad y^m + \alpha_1' y^{m-1} + \cdots + \alpha_m' = 0$$

in which at least one coefficient differs from the corresponding coefficient in (1). Subtracting (1) from (2), we get therefore a new equation of lower degree than m which is not identically zero and which is satisfied by every transform of x_1, say

$$(3) \qquad y^q + \beta_1 y^{q-1} + \cdots + \beta_q = 0.$$

If the β's are not all scalars, the degree can again be lowered by a repetition of this process, till finally an equation is reached with scalar coefficients not all zero; we can therefore regard the β's as scalars without loss of generality. Since however the identical equation is irreducible, the left-hand side of (3), with y replaced by ξ must be divisible by $\phi(\xi)$ whence it follows immediately that $\phi \equiv D$ i.e., $m = p$ and

$$\phi(\xi) = (\xi - x_p)(\xi - x_{p-1}) \cdots (\xi - x_1).$$

The linear factors are permutable cyclically since their product is a scalar.

The necessary modifications in the lemma when the algebra is not quadrate, will be obvious after the proof of theorem 1 below.

§ 3. Theorem I. *If its field, F, be suitably extended, any division algebra, A, can be expressed as the direct product of a commutative algebra, B, and a simple matric algebra. B is composed of all elements of A which are commutative with every other element, and its basis may be so chosen as to be rational in F.*

It has been shown elsewhere* that a division algebra, A, of order a, reduces to the direct sum of a number of simple matric algebras when the field is extended by the adjunction of a finite number of suitably chosen algebraic irrationalities. In this extended field, F', we may therefore write

$$A = A_1 + A_2 + \cdots + A_b,$$

where

$$A_i = (e_{pq}^{(i)}) \qquad (p, q = 1, 2, \cdots, a_i;\ \Sigma a_i^2 = a),$$

$$e_{pq}^{(i)} e_{qr}^{(i)} = e_{pr}^{(i)}, \qquad e_{pq}^{(i)} e_{rt}^{(i)} = 0 \quad (q \neq r), \qquad e_{pq}^{(i)} e_{rt}^{(j)} = 0 \qquad (i \neq j).$$

It is then obvious that the algebra, B, whose basis is $e_i = \sum_p e_{pp}^{(i)}$ $(i = 1, 2, \cdots, b)$ is composed of all elements of A which are commutative with every element of A. By expressing the basis of B in terms of any rational basis, any element y of B can be expressed in the form $y = \sum \xi_i x_i$, where the x's are rational elements of A, not all zero, and the ξ's are marks of F' which are linearly independent in F. If now x is any rational element, we have, from the definition of B, $xy = yx$; hence

$$0 = xy - yx = \sum \xi_i (xx_i - x_i x),$$

and therefore, since the ξ's are linearly independent in F, it follows that $xx_i - x_i x = 0$ for every x_i and x. The elements of the subalgebra generated by the elements x_i are therefore commutative with every element of A and this algebra, which is rational, is equivalent to B in F'.

If we extend the field F so as to include the elements of B, we get a new division algebra, A', of order a/b and when this field is again extended, A' reduces to a matric algebra which is simple, as otherwise B would not contain all elements commutative with every element of A. It follows immediately that all the algebras $A_1, A_2, \cdots, A_b$ have the same order a/b and that, in F', A is the direct product of a simple matric algebra C and the commutative algebra B.

If B reduces to the identity, A is said to be quadrate: its order is a square, and scalars are the only elements commutative with every element of the algebra.

THEOREM II. *If a division algebra, A, contains a quadrate subalgebra, B, it can be expressed as the direct product of B and another algebra C.*

If F' is the field F so extended as to render B reducible to the simple matric form, then it is known that A can be expressed as the direct product of B and an algebra C which contains all elements of A which are commutative with every element of B. Any element z of C can be expressed in the form $z = \sum \xi_s x_s$, where the x's are rational elements of A and the ξ's are marks

*Proceedings of the London Mathematical Society, Vol. 6 (1907), p. 102.

of F' which are linearly independent in F. If y is any element of B which is rational in F, then $yz = zy$; whence

$$0 = yz - zy = \sum \xi_s (yx_s - x_s y),$$

so that, as before, $yx_s = x_s y$ for every y of B. Hence the elements x_s belong to C which has therefore a basis rational in F.

Theorem III. *If A is a quadrate division algebra whose order, p^2, is the square of a prime, then the adjunction to the field F of any irrationality which renders any number of A wholly or partially reducible also reduces A to a matric algebra; or, in any field in which A is not primitive, it is equivalent to a matric algebra.*

This theorem follows immediately from theorem 1 and the theorem that any simple algebra is the direct product of a division algebra and a simple matric algebra, the latter being necessarily of order p^2 since p is a prime.

§ 4. The only types of division algebras hitherto discovered all come under the type described by L. E. Dickson.* These algebras are defined by the relations

$$xy = y\theta(x), \qquad y^n = g,$$

where the identical equation of x is a uniserial Abelian equation of degree n, and g is an element of the field of the coefficients which is not the norm of any rational polynomial in x. The following investigation serves to show the manner in which these algebras arise.

Let x be an element of a division algebra A in a field F, and let the order of the algebra, X, generated by x, be m, then by lemma 1 we can determine a complex $Z = (z_1, z_2, \cdots, z_n)$ such that $A = Xz_1 + Xz_2 + \cdots + Xz_n$, where A is of order mn. Every element y of A can therefore be expressed in the form $y = \sum g_r z_r$ where the g's are polynomials in x, and in particular we may set

$$\begin{aligned} z_1 x &= g_{11} z_1 + g_{12} z_2 + \cdots + g_{1n} z_n \\ z_2 x &= g_{21} z_1 + g_{22} z_2 + \cdots + g_{2n} z_n \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ z_n x &= g_{n1} z_1 + g_{n2} z_2 + \cdots + g_{nn} z_n . \end{aligned}$$

If G is the matrix (g_{pq}), whose coefficients lie in the field $F(x)$, these equations may be put in the form

$$(z_1, z_2, \cdots, z_n)x = G(z_1, z_2, \cdots, z_n),$$

and it follows immediately that

$$(z_1, z_2, \cdots, z_n)x^r = G^r(z_1, z_2, \cdots, z_n).$$

* Cf. Dickson, these Transactions, vol. 15 (1914), p. 31, and Wedderburn, ibid., p. 162.

Hence, if $m = n$, the matrix G satisfies the same identical equation as x and in any case G satisfies an equation of degree n whose coefficients may however contain x.*

In the field $F(x)$, x is always a root of this equation, the corresponding invariant axis being the modulus of the algebra. If now the identical equation of G is abelian, its roots are polynomials in x which are rational in F, say $\theta_r(x)$ $(r = 1, 2, \cdots, n;\ \theta_1(x) \equiv x)$, and to each root there corresponds a rational element of the algebra, say y_r, such that

$$y_r x = \theta_r(x) y_r;$$

and this leads to Dickson's algebra when the abelian equation is uniserial.†

§ 5. We shall now show that the Dickson algebra is the only quadrate division algebra of order 9.

Let x_1 be an element of such an algebra which is not commutative with any of its transforms so that its identical equation,

$$f(\xi) \equiv \xi^3 + a_1 \xi^2 + a_2 \xi + a_3 = 0, \tag{4}$$

is not abelian. By lemma 4, we can express $f(\xi)$ in the form

$$f(\xi) = (\xi - x_3)(\xi - x_2)(\xi - x_1)$$

where the factors may be permuted cyclically and no two of the x's are commutative.‡ Since $\xi - x_2$ is a right-hand factor of $f(\xi)$, and

$$x_2 = (x_2 - x_1) x_2 (x_2 - x_1)^{-1}$$

leads to $x_1 x_2 = x_2 x_1$ contrary to our assumption, therefore

$$x_3 = (x_2 x_1 - x_1 x_2) x_2 (x_2 x_1 - x_1 x_2)^{-1} \tag{5}$$

and by symmetry, permuting the suffixes cyclically, also

$$x_2 = (x_1 x_3 - x_3 x_1) x_1 (x_1 x_3 - x_3 x_1)^{-1},$$

$$x_1 = (x_3 x_2 - x_2 x_3) x_3 (x_3 x_2 - x_2 x_3)^{-1}.$$

But, from (4),

$$x_3 x_2 + x_2 x_1 + x_3 x_1 = a_2;$$

hence, permuting cyclically,

$$x_3 x_2 + x_2 x_1 + x_3 x_1 = x_2 x_1 + x_2 x_3 + x_1 x_3 = x_1 x_2 + x_1 x_3 + x_3 x_2,$$

* It can be proved in the same way that, to every element y_i of A, there corresponds a matrix Y_i, whose coefficients are scalar polynomials in x, such that $(z) y_i = Y_i(z)$; and if y_j is a second element of A and Y_j the corresponding matrix, the matrix belonging to $y_j y_i$ is $Y_i Y_j = (Y'_j Y'_i)'$.

† I have been unable to construct an algebra of this type which is not also a Dickson algebra i.e., one for which the equation is uniserial, but it appears probable that they exist.

‡ As the roots of $f(\xi)$ are necessarily distinct, any number commutative with x_1 is a scalar polynomial in x_1.

whence

$$x_2 x_1 - x_1 x_2 = x_1 x_3 - x_3 x_1 = x_3 x_2 - x_2 x_3 = y,$$

say, so that from (5)

$$x_1 = y x_3 y^{-1}, \qquad x_2 = y x_1 y^{-1}, \qquad x_3 = y x_2 y^{-1} = y^2 x_1 y^{-2}.$$

Therefore y^3 is commutative with x_1 and, as y is not a polynomial in x_1, it follows that $y^3 = h$ is a scalar. We may then assume that the identical equation of x_1 has the same form, say $x_1^3 = g$.

Let now $z_1 = x_1 y$, $z_2 = x_1 z_1 x_1^{-1} = x_1^2 y x_1^{-1}$, then

$$z_1 z_2 - z_2 z_1 = x_1 y x_1^2 y x_1^{-1} - x_1^2 y^2 = x_1 (y x_1^2 y - x_1 y^2 x_1) x_1^{-1};$$

but, since $x_2 x_1 x_3 = g$, we have $x_2 x_1 = x_3^2$, so that

$$0 = x_2 x_1 - x_3^2 = y x_1 y^{-1} x_1 - y^2 x_1^2 y^{-2} = y (x_1 y^2 x_1 - y x_1^2 y)/h,$$

since $y^3 = h$, and therefore $z_1 z_2 - z_2 z_1 = 0$, i.e., z_2 is a polynomial in z_1 so that the identical equation is a uniserial abelian equation and the algebra is of Dickson's type.

The identical equation of z_1 is easily found as follows:

$$z_1^2 = x_1 y x_1 y = y x_3 y x_3 = y^2 x_2 x_3 = y^2 x_3 x_2 - h,$$

therefore

$$z_1^3 + h z_1 = x_1 y \cdot y^2 x_3 x_2 = h x_1 x_3 x_2 = hg.$$

Princeton University

OSCILLATION THEOREMS FOR THE REAL, SELF-ADJOINT LINEAR SYSTEM OF THE SECOND ORDER[*]

BY

H. J. ETTLINGER

Introduction

It is the object of this paper to determine the number of oscillations of a linear combination of the form (2) for the systems (3) and (4). From these results, an oscillation theorem for the solution $u_p(x)$, corresponding to the pth characteristic number of (4), is obtained.

Given the second order self-adjoint linear differential equation

$$\frac{d}{dx}\left[K(x,\lambda)\frac{du}{dx}\right] - G(x,\lambda)u = 0 \tag{1}$$

and two linear combinations of a solution which does not vanish identically and its first derivative,

$$L_i[u(x,\lambda)] = \alpha_i(x,\lambda)u(x,\lambda) - \beta_i(x,\lambda)K(x,\lambda)u_x(x,\lambda) \quad (i=1,2), \tag{2}$$

we shall impose the following conditions and shall assume that they are satisfied throughout this paper:

I. $K(x,\lambda)$, $G(x,\lambda)$, $\alpha_i(x,\lambda)$, $\beta_i(x,\lambda)$, $\alpha_{ix}(x,\lambda)$, $\beta_{ix}(x,\lambda)$[†] are continuous, real functions of x in the interval

$$(X) \qquad (a \leqq x \leqq b)$$

and for all real values of λ in the interval

$$(\Lambda) \qquad (\mathcal{L}_1 < \lambda < \mathcal{L}_2).$$

II. $K(x,\lambda)$ is positive everywhere in (X,Λ) and

$$|\alpha_i| + |\beta_i| > 0$$

in (X,Λ).

III. For each value of x in (X), K and G decrease (or do not increase)

[*] Presented to the Society, Sept. 6, 1917.

[†] $f_{ix}(x,\lambda) = \frac{\partial}{\partial x}f_i(x,\lambda)$.

as λ increases. In no sub-interval of (X) are K and G simultaneously independent of λ and in no sub-interval of (X) is G identically zero.

IV. Either $\beta_i \equiv 0$ for all values of x and λ in (X, Λ); or else $\beta_i \neq 0$ in $a < x < b$ for all λ's in (Λ) and one of the following is true,

either (*a*) $\beta_i(a) \equiv 0$ for all λ's in (Λ), $\beta_i(b) \neq 0$, $-\alpha_i(b)/\beta_i(b)$ decreases (or does not increase) as λ increases;

or (*b*) $\beta_i(b) \equiv 0$, $\beta_i(a) \neq 0$ for all λ's in (Λ), $\alpha_i(a)/\beta_i(a)$ decreases (or does not increase) as λ increases;

or (*c*) $\beta_i(a) \neq 0$, $\beta_i(b) \neq 0$, $\alpha_i(a)/\beta_i(a)$ and $-\alpha_i(b)/\beta_i(b)$ decrease (or do not increase) as λ increases.

V.*

$$\lim_{\lambda \to \mathcal{L}_1} - \frac{\min\ G}{\min\ K} = -\infty,$$

$$\lim_{\lambda \to \mathcal{L}_2} - \frac{\max\ G}{\max\ K} = +\infty.$$

I. The Sturmian system

Concerning the system

$$\frac{d}{dx}\left[K(x, \lambda)\frac{du}{dx}\right] - G(x, \lambda)u = 0, \tag{3}$$

$$L_1[u(a, \lambda)] = 0, \qquad L_1[u(b, \lambda)] = 0,$$

Sturm's oscillation theorem† may be stated with Bôcher‡ substantially as follows:

The system (3) *satisfying conditions I–V has an infinite set of characteristic numbers such that*

$$\mathcal{L}_1 < \lambda_0 < \lambda_1 < \lambda_2 < \cdots < \mathcal{L}_2$$

and $U(x, \lambda_p)$, *the pth characteristic function, vanishes exactly* p *times on* $a < x < b$.

We seek to determine the number of oscillations of $L_1[U(x, \lambda_p)]$. We notice first that if $\beta_1 \equiv 0$ for all values of x and λ in (X, Λ), then all the zeros of $U(x, \lambda_p)$ and $L_1[U(x, \lambda_p)]$ coincide, since $\alpha_1(x, \lambda) \neq 0$. Hence we have for $\beta_1 \equiv 0$ precisely p zeros of $L_1[U(x, \lambda_p)]$ on $a < x < b$.

With Bôcher§ we define

$$\{\alpha_1 \beta_1\} = \beta_1 \alpha_{1x} - \alpha_1 \beta_{1x} + \frac{\alpha_1^2}{K} - \beta_1^2 G.$$

* This condition may be replaced by other sets of conditions. See Bôcher, *Leçons sur les Méthodes de Sturm* (hereafter referred to as *Leçons*) (1917), chap. III, paragraphs 13–15.

† Journal de Mathématiques pures et appliquées, vol. 1 (1836), p. 106 ff.

‡ *Leçons*, p. 63 ff.

§ *Leçons*, p. 45.

Let $\beta_1 \neq 0$ in (X, Λ), and consider the zeros of $U(x, \lambda)$ and $L_1[U(x, \lambda)]$. For $\lambda = \lambda_0$, $L_1[U(a, \lambda_0)] = 0$ and $L_1[U(b, \lambda_0)] = 0$, but $U(x, \lambda_0)$ does not vanish on (X). Hence $\{\alpha_1 \beta_1\}_{\lambda=\lambda_0}$ must vanish for some value of x in (X), since if $\{\alpha_1 \beta_1\}_{\lambda=\lambda_0} < 0$ throughout (X), $L_1[U(x, \lambda_0)]$ could vanish but once in (X);* and if $\{\alpha_1 \beta_1\}_{\lambda=\lambda_0} > 0$ for every x in (X), $U(x, \lambda_0)$ would vanish once in (X).† If $\{\alpha_1 \beta_1\} > 0$ for $\lambda \geqq \lambda_1$ for all values of x in (X), then the zeros of $L_1[U(x, \lambda)]$ and $U(x, \lambda)$ separate.† But $U(x, \lambda_p)$ vanishes exactly p times on $a < x < b$. Hence $L_1[U(x, \lambda_p)]$ vanishes $p - 1$ times on $a < x < b$ for $p \geqq 1$.

If $\beta_1(a, \lambda) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ for $a < x \leqq b$, then $\alpha(a) \neq 0$ in (Λ). Then the zeros of $L_1[U(x, \lambda_p)]$ and $U(x, \lambda_p)$ separate† each other on $a < x \leqq b$, provided $\{\alpha_1 \beta_1\}_{\lambda=\lambda_p} > 0$. Let x_1 be the first zero of $U(x, \lambda_p)$ on $a < x \leqq b$. Then

$$L_1[U(x_1, \lambda_p)] = -\beta_1(x_1, \lambda_p) K(x_1, \lambda_p) U_x(x_1, \lambda_p).$$

Now direct computation shows that

$$L_{1x}[U(a)] = \{\alpha_1 \beta_1\}_{x=a} \frac{K(a) U_x(a)}{\alpha_1(a)}.$$

But $K(a) U_x(a)$ and $K(x_1, \lambda_p) U_x(x_1, \lambda_p)$ have opposite signs, and $\operatorname{sgn} \beta_1(x_1, \lambda_p) = \operatorname{sgn} \beta_1(b)$.‡ Hence

$$\operatorname{sgn} L_1[U(x_1, \lambda_p)] = \operatorname{sgn} L_{1x}[U(a)] \cdot \operatorname{sgn} \beta_1(b) \operatorname{sgn} \alpha_1(a).$$

Thus $L_1[U(x, \lambda_p)]$ vanishes or does not vanish in $a < x < x_1$ according as $\alpha_1(a) \beta_1(b)$ is negative or is positive respectively. Hence if $\alpha_1(a) \beta_1(b) > 0$, $L_1[U(x, \lambda_p)]$ has $p - 1$ zeros on $a < x < b$, and if $\alpha_1(a) \beta_1(b) < 0$, $L_1[U(x, \lambda_p)]$ has p zeros on $a < x < b$, for $p \geqq 1$.

If $\beta_1(b, \lambda) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ for $a \leqq x < b$, then a similar argument can be made with the following result: if $\alpha_1(b) \beta_1(a) > 0$, $L_1[U(x, \lambda_p)]$ has $p - 1$ zeros on $a < x < b$, and if $\alpha_1(b) \beta_1(a) < 0$, $L_1[U(x, \lambda_p)]$ has p zeros on $a < x < b$, for $p \geqq 1$. Therefore

OSCILLATION THEOREM I. *If $U(x, \lambda_p)$ is the pth characteristic function of the system* (3) *satisfying conditions I–V and if $\{\alpha_1 \beta_1\} > 0$ for $\lambda \geqq \lambda_1$ for every x in (X), then $L_1[U(x, \lambda_p)]$ will vanish on $a < x < b$ for $p \geqq 1$,*
p times if
either $\beta_1 \equiv 0$ in (X, Λ);
or $\beta_1(a) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ in $a < x \leqq b$ and $\alpha_1(a) \beta_1(b) < 0$;
or $\beta_1(b) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ in $a \leqq x < b$ and $\alpha_1(b) \beta_1(a) < 0$;
$p - 1$ times if

* Bôcher, *Leçons*, p. 51.

† Ibid., p. 50.

‡ sgn f = sign of f.

either $\beta_1 \neq 0$ *in* (X, Λ);

or $\beta_1(a) \equiv 0$ *in* (Λ), $\beta_1(x, \lambda) \neq 0$ *in* $a < x \leqq b$ *and* $\alpha_1(a)\beta_1(b) > 0$;

or $\beta_1(b) \equiv 0$ *in* (Λ), $\beta_1(x, \lambda) \neq 0$ *in* $a \leqq x < b$ *and* $\alpha_1(b)\beta_1(a) > 0$.

Let $L[u(x, \lambda)] = \alpha(x, \lambda)u(x, \lambda) - \beta(x, \lambda)K(x, \lambda)u_x(x, \lambda)$, where $\alpha, \beta, \alpha_x, \beta_x$ are continuous in (X, Λ). Then we may state

OSCILLATION THEOREM II. *If* $U(x, \lambda_p)$ *is the pth characteristic function of* (3) *satisfying conditions I–V, where* $\{\alpha_1 \beta_1\} > 0$, $\{\alpha\beta\} > 0$ *for* $\lambda \geqq \lambda_1$ *and* $(\alpha_1 \beta - \alpha\beta_1) \neq 0$ *in* (X, Λ), *then* $L[U(x, \lambda_p)]$ *will vanish on* $a < x < b$, *for* $p \geqq 1$,

p times if

either $\beta_1 \neq 0$ *in* (X, Λ);

or $\beta_1(a) \equiv 0$ *in* (Λ), $\beta_1(x, \lambda) \neq 0$ *in* $a < x \leqq b$ *and* $\alpha_1(a)\beta_1(b) > 0$;

or $\beta_1(b) \equiv 0$ *in* (Λ), $\beta_1(x, \lambda) \neq 0$ *in* $a \leqq x < b$ *and* $\alpha_1(b)\beta_1(a) > 0$;

p + 1 times if

either $\beta_1 \equiv 0$ *in* (X, Λ);

or $\beta_1(a) \equiv 0$ *in* (Λ), $\beta_1(x, \lambda) \neq 0$ *in* $a < x \leqq b$ *and* $\alpha_1(a)\beta_1(b) < 0$;

or $\beta_1(b) \equiv 0$ *in* (Λ), $\beta_1(x, \lambda) \neq 0$ *in* $a \leqq x < b$ *and* $\alpha_1(b)\beta_1(a) < 0$.

Proof: The zeros of $L_1[U(x, \lambda_p)]$ and $L[U(x, \lambda_p)]$ separate one another on (X).* But $L_1[U(a, \lambda_p)] = 0$ and $L_1[U(b, \lambda_p)] = 0$. Hence $L[U(x, \lambda_p)]$ has one more zero on $a < x < b$, for $p \geqq 1$, than $L_1[U(x, \lambda_p)]$. This proves Theorem II.

II. The general self-adjoint system

Consider the system

$$
\frac{d}{dx}\left[K(x, \lambda)\frac{du}{dx}\right] - G(x, \lambda)u = 0,
$$

(4)

$$
L_1[u(a, \lambda)] = L_1[u(b, \lambda)], \qquad L_2[u(a, \lambda)] = L_2[u(b, \lambda)],
$$

where L_1 and L_2 are defined as in (2) and $\beta_2 \not\equiv 0$† in (X, Λ). We impose further conditions:

VI. $$\{\alpha_i \beta_i\} \neq 0 \text{ for } \lambda \geqq \lambda_1,$$

VII. $$\alpha_1 \beta_2 - \alpha_2 \beta_1 \equiv -1 \text{ in } (X, \Lambda),$$

VIII.‡ $$\begin{vmatrix} \alpha_1(a) & \beta_1(a) & \alpha_1(b) & \beta_1(b) \\ \alpha_2(a) & \beta_2(a) & \alpha_2(b) & \beta_2(b) \\ \Delta\alpha_1(a) & \Delta\beta_1(a) & \Delta\alpha_1(b) & \Delta\beta_1(b) \\ \Delta\alpha_2(a) & \Delta\beta_2(a) & \Delta\alpha_2(b) & \Delta\beta_2(b) \end{vmatrix} \geqq 0.$$

* Bôcher, *Leçons*, p. 50.

† This restriction does not involve a loss of generality, since if $\beta_1 \not\equiv 0$, $\beta_2 \equiv 0$ then we may interchange L_1 and L_2.

‡ $\Delta f = f(\lambda + \Delta\lambda) - f(\lambda)$ for $\Delta\lambda > 0$.

In a recent paper* the writer proved the following theorem concerning the system (4), satisfying conditions I–VIII:

There exists one and only one characteristic number of (4) *between every pair of characteristic numbers of the Sturmian system* (3). *If* λ_p *represents the ordered characteristic numbers of* (3) *and* l_p *those of* (4) (*account being taken of their multiplicity*) *then*

Case I. l_p *is in the interval* $(\lambda_p, \lambda_{p+1})$ *if* $L_2[U(b, \lambda_0)] \cdot \phi(\mathcal{L}_1 + \epsilon) > 0$.†

Case II. l_p *is in the interval* $(\lambda_{p-1}, \lambda_p)$, $p \geqq 1$, *if*

$$L_2[U(b, \lambda_0)] \cdot \phi(\mathcal{L}_1 + \epsilon) < 0.$$

Let $u_p(x) = u(x, l_p)$ be the pth characteristic function of (4). We proceed to consider the number of oscillations of $L_1[u_p(x)]$. We notice first that if $\lambda = l_p$ is a double characteristic number, then l_p coincides with λ_{p-1}, λ_p, or λ_{p+1}, and $L_1[u_p(x)]$ will have the number of oscillations designated by Theorem I.

If $\lambda = l_p$ is a simple value, we may discriminate between Case I and Case II, following a method due to Birkhoff.‡ If $\beta_1 \equiv 0$ in (X, Λ), the sign of $L_2[U(b, \lambda_0)]$ is the same as $-\alpha_1(a)\alpha_1(b)$, which is negative, since $\alpha_1 \neq 0$ in (X, Λ). Hence $L_2[U(b, \lambda_0)]$ is negative. Also $\phi(\mathcal{L}_1 + \epsilon)$ has the sign of $-\beta_2(a)\beta_2(b)$, but $\beta_2 \neq 0$ in (X, Λ). Hence $\phi(\mathcal{L}_1 + \epsilon)$ is negative, and we have Case I where l_p is on the interval $\lambda_p < \lambda < \lambda_{p+1}$. By Theorem I, $L_1[U(x, \lambda_p)]$ vanishes exactly p times on $a < x < b$ for $p \geqq 1$, and $L_1[U(x, \lambda_{p+1})]$ vanishes $p + 1$ times on $a < x < b$ for $p \geqq 1$. But $L_1[U(b, \lambda)] \neq 0$ for $\lambda_p < \lambda < \lambda_{p+1}$. Hence $L_1[U(x, \lambda)]$ vanishes $p + 1$ times on $a < x < b$ for $\lambda_p < \lambda < \lambda_{p+1}$. But the roots of $L_1[U(x, \lambda)]$ and $L_1[u(x, \lambda)]$§ separate one another, and $L_1[U(a, \lambda)] = 0$. Hence $L_1[u(x, \lambda)]$ vanishes $p + 1$ or $p + 2$ times on $a < x < b$ for $\lambda_p < \lambda < \lambda_{p+1}$, $p \geqq 1$. Hence $L_1[u_p(x)]$ vanishes either $p + 1$ or $p + 2$ times on $a < x < b$, $p \geqq 1$. But from (4) the number of zeros of $L_1[u_p(x)]$ is always even. Therefore we have $p + 2$ roots if p is even and $p + 1$ roots for p odd, $p \geqq 1$.

If $\beta_1(a) \equiv 0$ in (Λ) and $\beta_1(b) \neq 0$, the sign of $L_2[U(b, \lambda_0)]$ is that of $\alpha_1(a)/\beta_1(b)$, and $\phi(\mathcal{L}_1 + \epsilon)$ has the sign of $\beta_1(b)/\alpha_1(a)$. Hence we have Case I where l_p is on the interval $\lambda_p < \lambda < \lambda_{p+1}$. If $\alpha_1(a)\beta_1(b) < 0$, by Theorem I, $L_1[U(x, \lambda_p)]$ vanishes p times on $a < x < b$ and $L_1[U(x, \lambda_{p+1})]$ vanishes $p + 1$ times on $a < x < b$. Reasoning exactly as above, we find that $L_1[u_p(x)]$ vanishes $p + 2$ times on $a < x < b$ if p is even and $p + 1$

* *Existence Theorems for the General, Real, Self-Adjoint Linear System of the Second Order*, these Transactions, vol. 19 (1918), p. 94.

† $\phi(\lambda) = 0$ is the characteristic equation of (4) whose roots are the characteristic numbers, l_p. $L_1 + \epsilon$ is a value of λ in (Λ) near L_1.

‡ *Existence and Oscillation Theorem for a Certain Boundary Value Problem*, these Transactions, vol. 10 (1909), pp. 259–270.

§ Bôcher, *Leçons*, p. 48.

times if p is odd, $p \geqq 1$. If $\alpha_1(a)\beta_1(b) > 0$, by Theorem I, $L_1[U(x, \lambda_p)]$ vanishes $p - 1$ times on $a < x < b$, and $L_1[U(x, \lambda_{p+1})]$ vanishes p times on $a < x < b$. But $L_1[U(b, \lambda)] \neq 0$ for $\lambda_p < \lambda < \lambda_{p+1}$, and $L_1[U(a, \lambda)] = 0$. The roots of $L_1[U(x, \lambda)]$ and $L_1[u(x, \lambda)]$ separate one another. Hence $L_1[u_p(x)]$ vanishes either p or $p + 1$ times on $a < x < b$, $p \geqq 1$. But from (4) the number of zeros of $L_1[u_p(x)]$ is always even. Therefore we have p roots if p is even and $p + 1$ roots if p is odd.

If $\beta_1(b) \equiv 0$ in (Λ) and $\beta_1(a) \neq 0$, the sign of $L_2[U(b, \lambda_0)]$ is that of $-\beta_1(a)/\alpha_1(b)$, and $\phi(\mathcal{L}_1 + \epsilon)$ has the sign of $-\beta_1(a)/\alpha_1(b)$. Hence we have Case I again. If $\alpha_1(b)\beta_1(a) < 0$ we proceed as before and obtain $p + 2$ zeros of $L_1[u_p(x)]$ on $a < x < b$ if p is even and $p + 1$ zeros if p is odd, $p \geqq 1$. If $\alpha_1(b)\beta_1(a) > 0$, we obtain p zeros of $L_1[u_p(x)]$ on $a < x < b$ if p is even and $p + 1$ zeros if p is odd, $p \geqq 1$.

If $\beta_1 \neq 0$ in (X, Λ), the sign of $L_2[U(b, \lambda_0)]$ is that of $\beta_1(a)/\beta_1(b)$, which is positive. The sign of $\phi(\mathcal{L}_1 + \epsilon)$ is that of $\beta_1(a)\beta_2(b) - \beta_2(a)\beta_1(b)$ if $\beta_1(a)\beta_2(b) - \beta_2(a)\beta_1(b)$ does not vanish. If

$$\beta_1(a)\beta_2(b) - \beta_2(a)\beta_1(b) = 0,$$

the sign of $\phi(\mathcal{L}_1 + \epsilon)$ is that of $\beta_1(a)\beta_1(b)$, which is positive. Accordingly, if $\beta_1(a)\beta_2(b) - \beta_2(a)\beta_1(b) \geqq 0$, we have Case I, and l_p is on $\lambda_p < \lambda < \lambda_{p+1}$. By Theorem I, $L_1[U(x, \lambda_p)]$ vanishes $p - 1$ times, and, proceeding as above, we find that $L_1[u_p(x)]$ has p zeros if p is even and $p + 1$ zeros if p is odd, $p \geqq 1$. Likewise, if $\beta_1(a)\beta_2(b) - \beta_2(a)\beta_1(b) < 0$, l_p is on $\lambda_{p-1} < \lambda < \lambda_p$, and $L[u_p(x)]$ will vanish p times if p is even and $p - 1$ times if p is odd, $p \geqq 1$.

Summarizing the various kinds of coefficients of the boundary conditions of (4) as follows:

A: $\beta_1 \equiv 0$ in (X, Λ);

B^-: Either $\beta_1(a) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ in $a < x \leqq b$, and $\alpha_1(a)\beta_1(b) < 0$;
or $\beta_1(b) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ in $a \leqq x < b$, and $\alpha_1(b)\beta_1(a) < 0$;

B^+: Either $\beta_1(a) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ in $a < x \leqq b$, and $\alpha_1(a)\beta_1(b) > 0$;
or $\beta_1(b) \equiv 0$ in (Λ), $\beta_1(x, \lambda) \neq 0$ in $a \leqq x < b$, and $\alpha_1(b)\beta_1(a) > 0$;

C^+: $\beta_1 \neq 0$ in (X, Λ) and $\beta_1(a)\beta_2(b) - \beta_2(a)\beta_1(b) \geqq 0$ in (Λ);

C^-: $\beta_1 \neq 0$ in (X, Λ) and $\beta_1(a)\beta_2(b) - \beta_2(a)\beta_1(b) < 0$ in (Λ);

we state

OSCILLATION THEOREM III. *If $u_p(x)$ is the pth characteristic function corresponding to a simple value of the system* (4) *satisfying conditions I–VIII,* *then the number of zeros of $L_1[u_p(x)]$ on $a < x < b$ for $p \geqq 1$ is given by the following table:*

Case	*Number of zeros*	
	$p = 2m$	$p = 2m + 1$
A *or* B^-	$p + 2$	$p + 1$
B^+ *or* C^-	p	$p + 1$
C^+	p	$p - 1$.

Furthermore, we notice that, since $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$ in (X, Λ), the zeros of $L_1[u_p(x)]$ and $L_2[u_p(x)]$ separate one another on X,* and, if $\lambda = l_p$ is a double value, the number of zeros of $L_2[u_p(x)]$ is given by Theorem II. If $\lambda = l_p$ is a simple value, both $L_1[u_p(x)]$ and $L_2[u_p(x)]$ by (4) have an even number of zeros on (X). Hence they oscillate the same number of times. Therefore

COROLLARY.. *The number of zeros of $L_2[u_p(x)]$ on $a < x < b$, for $p \geqq 1$, is precisely the number given in the table of Theorem III.*

From the foregoing results we may deduce the number of oscillations of $u_p(x)$. If $\lambda = l_p$ is a double value, then $u_p(x)$ will differ from $U(x, \lambda_{p-1})$, $U(x, \lambda_p)$, or $U(x, \lambda_{p+1})$ by at most a non-vanishing factor, and the number of zeros of $u_p(x)$ will be given by the Sturmian Oscillation Theorem.

For a simple value we consider the following cases:

If $\beta_1 \equiv 0$ in (X, Λ) the zeros of $u_p(x)$ and $L_1[u_p(x)]$ coincide. Hence $u_p(x)$ has $p + 2$ zeros if $p = 2m$ and $p + 1$ zeros if $p = 2m + 1$.

If $\beta_1 \neq 0$ in (X, Λ) or $\beta_1(a) \equiv 0$, $\beta_1(b) \neq 0$ or $\beta_1(b) \equiv 0$, $\beta_1(a) \neq 0$, we write the first boundary condition of (4)

$$u_p(a) \cdot P(a) = u_p(b) \cdot P(b),$$

where

$$P_1(x) = \alpha_1(x) - \beta_1(x)K(x)u_p'(x)/u_p(x).$$

If $P_1(a)P_1(b) > 0$, $u_p(a)u_p(b)$ will be positive and $u_p(x)$ has an even number of roots. But the zeros of $u_p(x)$ and $L_1[u_p(x)]$ separate. Hence $u_p(x)$ and $L_1[u_p(x)]$ have the same number of zeros on $a < x < b$. If $P_1(a)P_1(b) < 0$, $u_p(a)u_p(b)$ will be negative and $u_p(x)$ has an odd number of zeros. Hence $u_p(x)$ will have one more or one less zero than $L_1[u_p(x)]$ on $a < x < b$.

It is also to be noticed that in Case I, since l_p is on $(\lambda_p, \lambda_{p+1})$, $u_p(x)$ can vanish not more than $p + 2$ times nor less than $p - 1$ times. In Case II, since l_p is on $(\lambda_{p-1}, \lambda_p)$, $u_p(x)$ can vanish not more than $p + 1$ times nor less than $p - 2$ times.†

Designating the condition $P_1(a) \cdot P_1(b) > 0$ by P^+ and the condition $P_1(a) \cdot P_1(b) < 0$ by P^-, we may state

* Bôcher, *Leçons*, p. 50.

† Cf. Cor., p. 96 of the paper of the author referred to above.

OSCILLATION THEOREM IV. *If $u_p(x)$ is the pth characteristic function corresponding to a simple value of the system* (4) *satisfying conditions I–VIII, then the number of zeros of $u_p(x)$ on $a < x < b$, for $p \geqq 1$, is given by the following table:*

Case	*Number of zeros*	
	$p = 2m$	$p = 2m + 1$
A	$p + 2$	$p + 1$
$B^- P^+$	$p + 2$	$p + 1$
$B^- P^-$	$p + 1$	$p + 2$ *or* p
$B^+ P^+$	p	$p + 1$
$B^+ P^-$	$p + 1$ *or* $p - 1$	$p + 2$ *or* p
$C^+ P^+$	p	$p + 1$
$C^+ P^-$	$p + 1$ *or* $p - 1$	$p + 2$ *or* p
$C^- P^+$	p	$p - 1$
$C^- P^-$	$p + 1$ *or* $p - 1$	p *or* $p - 2$

Note: if in particular we choose

$$\bar{\alpha}_i(x, \lambda) = \frac{(b - x)\alpha_i(\lambda) + (x - a)\gamma_i(\lambda)}{b - a}$$

$$\bar{\beta}_i(x, \lambda) = \frac{(b - x)\beta_i(\lambda) + (a - x)\delta_i(\lambda)}{b - a}$$

the system (4) becomes identical with that of (4) of the paper by the writer to which reference has been made above. If $\beta_i \delta_i > 0$, it will be necessary to modify the boundary conditions of (4) by taking

$$L_i[u(a, \lambda)] = - L_i[u(b, \lambda)]$$

with condition VII replaced by

$$\alpha_1(a)\beta_2(a) - \alpha_2(a)\beta_1(a) = \alpha_2(b)\beta_1(b) - \alpha_1(b)\beta_2(b) = -1,$$

if only one of the boundary conditions is modified. If both boundary conditions are modified, condition VII remains unchanged. In either case conditions I–VI and VIII will be satisfied. The oscillation theorems will be true with the modification that $L_i[u_p(x)]$ will have an odd number of zeros.

UNIVERSITY OF TEXAS,
AUSTIN, TEXAS

NEW PROOFS OF CERTAIN FINITENESS THEOREMS IN THE THEORY OF MODULAR COVARIANTS*

BY

OLIVE C. HAZLETT

1. Introduction. In a recent paper of mine† it was proved that every modular covariant of a system of forms S (with variables x and y) is a polynomial in the universal covariant L and modular invariants of the system of forms S (with variables ξ and η) enlarged by the linear form $\eta x - \xi y$ which have been made formally invariant as to x and y. As pointed out in that paper, we have as a corollary the following:

If K is the class of all modular concomitants of the system S which are formally invariant as to certain sets of coefficients and variables, but not formally invariant as to x and y, then the theorem tells us how to construct the set K' of all modular concomitants which are formally invariant as to x and y in addition to being formally invariant as to those sets of coefficients and variables with respect to which K is formally invariant.

Here x and y may be the variables of the system S or a pair of variables which is cogredient with the variables of the system or even a pair of variables which is cogredient with the variables aside from a power of the determinant of the transformation. In fact, every modular covariant of the set K' is a polynomial in L and the concomitants of the set K which have been made formally invariant as to x and y. In the present paper we give a few extensions and applications of this theorem.

§§ 2–3. New proofs of some finiteness theorems

2. Finiteness theorem for modular covariants. Dickson has already shown that modular invariants from their very nature have the finiteness property.‡ He has also proved§ that modular covariants have the finiteness property. We now proceed to give a proof of the finiteness theorem for modular covariants directly from the finiteness theorem for modular invariants.

* Presented to the Society, October 25, 1919.

† *A Theorem on Modular Covariants*, these Transactions, vol. 21 (1920), p. 247.

‡ *General Theory of Modular Invariants*, these Transactions, vol. 10 (1909), p. 126.

§ *Proof of the Finiteness of Modular Covariants*, these Transactions, vol. 14 (1913), p. 300.

For, in the first place, every modular covariant of a system S of forms is a polynomial in L and the modular invariants of the enlarged system S' which have been made formally invariant as to x and y. Now the modular invariants of S' have the "finiteness" property—i.e., they are all expressible as polynomials in a finite subset. This is not, however, the same as saying that the modular invariants of S' which have been made formally invariant as to x and y have the finiteness property, since any one modular invariant of S' produces a number of modular invariants which are formally invariant as to x and y. Let $I_1, I_2, \cdots, I_\nu$ be modular invariants of S' which are congruent to I when x and y are in the field and are formally invariant as to x and y. Moreover, let I_1 be such an invariant of the lowest possible degree, say d, in x and y; let I_2 be of degree $d + p^n - 1$, if there be any such;* let I_3 be of degree $d + 2(p^n - 1)$, if there be any such; and finally let I_ν be of degree $d + (p^n - 1)^2$. From the proof of the fundamental theorem of the paper mentioned in the introduction, every invariant of S' of degree d which is formally invariant as to x and y and which is congruent to I when x and y are in the field is of the form $I_1 + LI_1$, where I_1 is a modular invariant of S' which is formally invariant as to x and y and is of a degree in x and y which is less than d. A similar remark applies to invariants of S' which are congruent to I when x and y are in the field and which are of degree $d + (p^n - 1), \cdots, d + (p^n - 1)^2$ in x and y.

We can readily construct an invariant of S' which is of degree $d + p^n(p^n - 1)$ and is congruent to I when x and y are in the field, for QI_1 is such an invariant. Similarly, we can express every modular invariant of S' which is formally invariant as to x and y and which is congruent to I whenever x and y are in the field as a polynomial in $I_1, I_2, \cdots, I_\nu, L, Q$ and modular invariants of lower order which are formally invariant as to x and y. Hence, by induction, we prove the finiteness theorem for the modular invariants of S' which are formally invariant as to x and y.

In fact, we can so choose a fundamental set of invariants of S' which are formally invariant as to x and y that to each invariant of a fundamental set of modular invariants of S there correspond at most p^n invariants in a fundamental set of invariants of S' which are formally invariant as to x and y. In addition, we have to put L and Q in our fundamental set. We can, if we wish, think of Q as a modular invariant of S' which is formally invariant as to x and y, arising from the modular invariant 1. Notice that the degree of 1 is congruent to $p^n(p^n - 1)$ modulo $p^n - 1$. Also, L can be regarded as a

* It is to be noted that there is not necessarily any such invariant of degree $d + p^n - 1$, nor of degree $d + 2(p^n - 1)$, etc. For example, if I be unity, the only homogeneous formal invariants congruent to I for values of x and y in the field are 1 and powers of $Q = I_1$. In this case there are no invariants $I_2, I_3, \cdots, I_\nu$.

modular invariant of S' which has been made formally invariant as to x and y—it is a formal invariant arising from the modular invariant 0. At first there appears to be a discrepancy here, since the degree of L is $p^n + 1 \not\equiv 0 \pmod{p^n - 1}$. This discrepancy, however, is only apparent; for, if x and y are in the field, L becomes $xy - xy = 0$. In other words, L reduces to a quadratic which "telescopes," so to speak.

Hence, applying the fundamental theorem quoted in the introduction, we have proved

THEOREM I. *The set of all modular covariants of a system S of binary forms has the finiteness property—i.e., there is a finite number of covariants of the set such that every covariant of S is expressible as a polynomial in the covariants of the subset.*

3. Finiteness theorem for modular invariants of a system of forms and cogredient points. The above theorem may, for convenience, be restated thus: Let K be the class of all modular concomitants of the system Σ which are formally invariant as to certain sets of coefficients and variables, but not formally invariant as to x and y; and let K' be the class of all modular concomitants of the same system Σ which are formally invariant as to x and y in addition to being formally invariant as to those sets of coefficients and variables with respect to which K is formally invariant. Here, as in § 1, x and y may be the variables of the system Σ or a pair of variables which is cogredient with the variables of the system or even a pair of variables which is cogredient with the variables aside from a power of the determinant of the transformation. Then, if the set K has the finiteness property, the set K' has the finiteness property.

Now consider the set of all modular invariants of a system of forms S and the cogredient points $(x_1, y_1), (x_2, y_2), \cdots, (x_k, y_k)$. In the body of the proof of the fundamental theorem of the Chicago paper,* it was proved that the set of all invariants of the system S and the cogredient points is identical with the set of all invariants of the system S enlarged by the linear forms $\eta_1 x_1 - \xi_1 y_1, \eta_2 x_2 - \xi_2 y_2, \cdots, \eta_k x_k - \xi_k y_k$ where it is understood that now the ξ's and η's are the variables and the x's and y's are coefficients which are independent variables. That is, the invariants of S and the cogredient points are the invariants of the enlarged system S' which are formally invariant as to the x's and y's.

Let S' be the system Σ of the beginning of this section, and apply Theorem I as reworded above, making the set of modular invariants of S' formally invariant as to the pairs (x_i, y_i) one at a time. By induction, we thus prove

THEOREM II. *The set of all modular invariants of a system S of binary forms*

* These Transactions, vol. 21 (1920), pp. 251–252 and p. 254.

and the cogredient points (x_1, y_1), (x_2, y_2), $\cdots$, (x_k, y_k) *has the finiteness property.*

This is the theorem of Professor F. B. Wiley's Chicago dissertation;* but the present proof has the advantage of showing the relation between the modular invariants of S' on the one hand and the invariants of the original system S on the other hand. It also shows the relation between the invariants of S and $k - 1$ cogredient points, and the invariants of S and k cogredient points.

This theorem includes as a special case the finiteness theorem for the modular invariants of any number m of cogredient (binary) points. For $m = 1$, Dickson† has already shown that a fundamental set of invariants consists of L and Q. This affords a simple illustration of the fundamental theorem quoted in § 1. For the invariants of the point (x, y) are the same as the formal invariants of the linear form $l = \eta x - \xi y$, where ξ and η are the variables, and hence (by the theorem) are simply the modular invariants of l which have been made formally invariant as to x and y. Now there are two classes of linear forms l—(i) when $x = y = 0$, (ii) when x and y are in the field but $(x, y) \neq (0, 0)$. Accordingly we may take as a set of modular invariants which characterize the classes for l: (i) a function which is $= 0$ whenever x and y are in the field, (ii) a function which is $= 1$ when x and y are in the field, but $(x, y) \neq (0, 0)$, and $= 0$ when $(x, y) = (0, 0)$. By Dickson's fundamental memoir‡ all modular invariants are linear combinations of these two functions. Now L is a formal invariant which satisfies condition (i), and Q is a formal invariant which satisfies condition (ii).

If $m = 2$, a fundamental set of invariants§ for the Galois field $GF[p]$ is

$$L_i = \begin{vmatrix} x_i^p & y_i^p \\ x_i & y_i \end{vmatrix}, \qquad Q_i = \frac{1}{L_i}\begin{vmatrix} x_i^{p^2} & y_i^{p^2} \\ x_i & y_i \end{vmatrix} \qquad (i = 1, 2),$$

$$M = x_2 y_1 - y_2 x_1, \qquad M_1 = x_2 y_1^p - y_2 x_1^p, \qquad M_2 = x_2^p y_1 - y_2^p x_1,$$

$$N_s = \frac{M_2^{s+1} L_1^{p-s-1} + (-1)^s M_1^{p-s} L_2^s}{M^p} \qquad (1 \leq s \leq p - 2).$$

The invariants Q_i and N_s are all integral functions of x_1, x_2, y_1, y_2. This gives another simple illustration of the theorem of my Chicago paper. For the invariants of the two points (x_1, y_1) and (x_2, y_2) are identical with the formal invariants of two linear forms $l_1 = \eta_1 x_1 - \xi_1 y$ and $l_2 = \eta_2 x_2 - \xi_2 y_2$ (where the variables are the ξ's and η's), which in turn are the modular

[*] These Transactions, vol. 15 (1914), pp. 431–438.

[†] *Madison Colloquium Lectures*, p. 38; these Transactions, vol. 12 (1911), p. 1; Quarterly Journal of Mathematics, 1911, p. 158.

[‡] These Transactions, vol. 10 (1909), p. 126.

[§] W. C. Krathwohl, *Modular invariants of two pairs of cogredient variables* (Chicago dissertation), American Journal of Mathematics, vol. 36 (1914), pp. 449–460.

invariants of l_1 and l_2 which have been made formally invariant as to the x's and y's. Now the classes of the pairs of linear forms are

(1) $(x_1, y_1) = (x_2, y_2) = (0, 0)$,

(2) $(x_1, y_1) = (0, 0)$, (x_2, y_2) in the field but $\neq (0, 0)$,

(3) $(x_2, y_2) = (0, 0)$, (x_1, y_1) in the field but $\neq (0, 0)$,

$(4)_\lambda$ $(x_1, y_1) = (a, b)$, $(x_2, y_2) = (\lambda a, \lambda b)$, $\lambda \neq 0$ and $(a, b) \neq (0, 0)$,

$(5)_\Delta$ $(x_1, y_1) = (a, b)$, $(x_2, y_2) = (c, d)$, where a, b, c, d are in the field, but $\Delta = ad - bc \neq 0$.

Now for the different classes, the invariants above take the following values

	L_1, L_2	Q_1	Q_2	M, M_1, M_2	N_s
Class 1	0	0	0	0	0
Class 2	0	0	1	0	0
Class 3	0	1	0	0	0
Class 4_λ	0	1	1	0	λ^{ps}
Class 5_Δ	0	1	1	Δ	0

Thus the theorem is verified for this special case.

§§ 4–9. Application to formal covariants

4. A lemma. In this section we will prove an important lemma which shows the intimate relation between modular covariants and formal covariants. This lemma is not new, but is a special case of Miss Sanderson's theorem.* The present proof is given because it is elementary in nature and because it furnishes a simple formula for a formal covariant which is congruent to a given modular covariant C whenever the coefficients are in the field.

Let $\phi(a, b, c, \cdots; x, y) = C$ be a modular covariant of the system S under the group G of linear transformations with coefficients in the field $GF[p^n]$. Moreover, let ϕ be of index w. There is no loss of generality in assuming that ϕ is pseudo-homogeneous† in the k coefficients $a, b, c, \cdots$ of degree d. For, if ϕ is not pseudo-homogeneous in the coefficients, it is the sum of a finite number of modular covariants which are pseudo-homogeneous in the $a, b, c, \cdots$.

First, we construct a function K which is homogeneous in the k independent

* These Transactions, vol. 14 (1913), pp. 489–500.

† A function f is said to be pseudo-homogeneous of degree d if, when the arguments $a, b, c, \cdots$ are multiplied by ρ, any non-zero mark of the field $GF[p^n]$, the function f is multiplied by ρ^d. If f is a polynomial, this means that the degrees of the different terms of f differ at most by integral multiples of $p^n - 1$.

variables $a, b, c, \cdots$ and such that $K \equiv C$ whenever $a, b, c, \cdots$ are in the field. We can take

$$K = \sum \left[\phi(a_0, b_0, c_0, \cdots; x, y) \left\{ \frac{Q\begin{pmatrix} a, & b, & c, & \cdots \\ a_0, & b_0, & c_0, & \cdots \end{pmatrix}}{Q\begin{pmatrix} a_0, & b_0, & c_0, & \cdots \\ a_0, & b_0, & c_0, & \cdots \end{pmatrix}} \right\}^d \right].$$

Here Σ denotes the sum of all terms of the type indicated as the k-tuple $(a_0, b_0, c_0, \cdots)$ ranges over the k-tuples of a certain set σ defined below. Inside the bracket,

$$Q\begin{pmatrix} a, & b, & c, & \cdots \\ a_0, & b_0, & c_0, & \cdots \end{pmatrix}$$

stands for

$$\frac{\begin{vmatrix} a & b & c & \cdots \\ a^{p^n} & b^{p^n} & c^{p^n} & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ a^{p^{(k-2)n}} & b^{p^{(k-2)n}} & c^{p^{(k-2)n}} & \cdots \\ a^{p^{(k-1)n}} & b^{p^{(k-1)n}} & c^{p^{(k-1)n}} & \cdots \end{vmatrix}}{\begin{vmatrix} a & b & c & \cdots \\ a^{p^n} & b^{p^n} & c^{p^n} & \cdots \\ \cdot & \cdot & \cdot & \cdot \\ a^{p^{(k-2)n}} & b^{p^{(k-2)n}} & c^{p^{(k-2)n}} & \cdots \\ a_0 & b_0 & c_0 & \cdots \end{vmatrix}}$$

and

$$Q\begin{pmatrix} a_0, & b_0, & c_0, & \cdots \\ a_0, & b_0, & c_0, & \cdots \end{pmatrix}$$

stands for the value of Q when $a = a_0, b = b_0, c = c_0, \cdots$. Miss Sanderson has already shown that $Q \equiv 0$ if and only if $a, b, c, \cdots$ are not proportional to $a_0, b_0, c_0, \cdots$.* The numerator of Q is the product of all essentially distinct linear functions of $a, b, c, \cdots$ in which the coefficients are marks of the field, and the denominator of Q is the product of all those essentially distinct linear functions of $a, b, c, \cdots$ (in which the coefficients are marks of the field) which vanish whenever the $a, b, c, \cdots$ are proportional to $a_0, b_0, c_0, \cdots$. Notice that $Q\begin{pmatrix} a, & b, & c, & \cdots \\ a_0, & b_0, & c_0, & \cdots \end{pmatrix}$ is a polynomial of degree $p^{(k-1)n}$ in the coefficients $a, b, c, \cdots$, and thus K is homogeneous in the $a, b, c, \cdots$ of degree $d(p^{(k-1)n})$ which is congruent to d modulo $p^n - 1$.†

Into the set σ we put k-tuples $(a_i, b_i, c_i, \cdots)$—where the $a_i, b_i, c_i, \cdots$ are in the field—such that, if any particular k-tuple $(a_i, b_i, c_i, \cdots)$ is in the set σ,

* These Transactions, vol. 14 (1913), p. 491.

† These Transactions, vol. 14 (1913), pp. 492, 493.

then $(\rho a_i, \rho b_i, \rho c_i, \cdots)$ is not in the set σ when ρ is any non-zero mark of the field; and such, moreover, that if $(\alpha, \beta, \gamma, \cdots)$ is any k-tuple of marks of the field, then there is in the set σ some k-tuple $(a_i, b_i, c_i, \cdots)$ such that $\alpha = \rho a_i, \beta = \rho b_i, \gamma = \rho c_i, \cdots$ for ρ some non-zero mark of the field.

If we now give to $a, b, c, \cdots$ any set of values in the field, this set of values is of the form $\rho a_1, \rho b_1, \rho c_1, \cdots$, where $(a_1, b_1, c_1, \cdots)$ is a k-tuple of the set σ. Then K has the value

$$\sum \left[\phi(a_0, b_0, c_0, \cdots; x, y) \left\{ \frac{Q\begin{pmatrix} a_1, b_1, c_1, \cdots \\ a_0, b_0, c_0, \cdots \end{pmatrix}}{Q\begin{pmatrix} a_0, b_0, c_0, \cdots \\ a_0, b_0, c_0, \cdots \end{pmatrix}} \right\}^d \rho^{dp^{(k-1)n}} \right]$$

$$\equiv \phi(a_1, b_1, c_1, \cdots; x, y)\rho^d$$

$$\equiv \phi(\rho a_1, \rho b_1, \rho c_1, \cdots; x, y) \equiv C$$

since C is pseudo-homogeneous in $a, b, c, \cdots$ of degree d.

If we now subject the variables x and y to any transformation of determinant Δ of the group G, the indeterminates $a, b, c, \cdots$ are subjected to an induced transformation which carries any particular k-tuple of the set σ into a k-tuple of the form $(\rho a_1, \rho b_1, \rho c_1, \cdots)$ where $(a_1, b_1, c_1, \cdots)$ is a k-tuple of the set σ and ρ is some mark of the field. Let the indeterminates $a, b, c, \cdots$ go into $a', b', c', \cdots$ and let $a_0, b_0, c_0, \cdots$ go into $\rho a_0', \rho b_0', \rho c_0', \cdots$ where $(a_0', b_0', c_0', \cdots)$ is a k-tuple of the set σ. Since C is a modular covariant of index w which is pseudo-homogeneous in $a, b, c, \cdots$ of degree d,

$$\Delta^w \phi(a_0, b_0, c_0, \cdots; x, y) = \rho^d \phi(a_0', b_0', c_0', \cdots; x', y').$$

At the same time, since $(a, b, c, \cdots)$, $(a^{p^n}, b^{p^n}, c^{p^n}, \cdots)$, etc. are cogredient,

$$Q\begin{pmatrix} a, & b, & c, & \cdots \\ a_0, & b_0, & c_0, & \cdots \end{pmatrix} \equiv \frac{1}{\rho} Q\begin{pmatrix} a', b', c', \cdots \\ a_0', b_0', c_0', \cdots \end{pmatrix}$$

and

$$Q\begin{pmatrix} a_0, b_0, c_0, \cdots \\ a_0, b_0, c_0, \cdots \end{pmatrix} \equiv \rho^{p^{(k-1)n}-1} Q\begin{pmatrix} a_0', b_0', c_0', \cdots \\ a_0', b_0', c_0', \cdots \end{pmatrix}.$$

Thus

$$\Delta^w K = \sum \left[\rho^d \phi(a_0', b_0', c_0', \cdots; x', y') \left\{ \frac{Q\begin{pmatrix} a', b', c', \cdots \\ a_0', b_0', c_0', \cdots \end{pmatrix}}{Q\begin{pmatrix} a_0', b_0', c_0', \cdots \\ a_0', b_0', c_0', \cdots \end{pmatrix}} \right\}^d \frac{1}{\rho^{dp^{(k-1)n}}} \right]$$

$$\equiv \sum \left[\phi(a_0', b_0', c_0', \cdots; x', y') \left\{ \frac{Q\begin{pmatrix} a', b', c', \cdots \\ a_0', b_0', c_0', \cdots \end{pmatrix}}{Q\begin{pmatrix} a_0', b_0', c_0', \cdots \\ a_0', b_0', c_0', \cdots \end{pmatrix}} \right\}^d \right] = K'.$$

Notice that in this congruence the $a, b, c, \cdots$ and the $a', b', c', \cdots$ are two sets of indeterminates. Thus the lemma is proved.

5. A theorem on formal covariants. Let C be a modular covariant of the system S of forms with coefficients $a, b, c, \cdots$. As above, there is no loss of generality in assuming that C is pseudo-homogeneous of degree d in the coefficients as well as homogeneous in the variables x and y. By the preceding section, we know that there is at least one homogeneous formal covariant K of the system S which is congruent to C whenever the coefficients $a, b, c, \cdots$ are marks of the field. Let K_0 be one such covariant of lowest degree ω in $a, b, c, \cdots$. If there is a second covariant K which is of degree ω in $a, b, c, \cdots$ and which is congruent to C whenever $a, b, c, \cdots$ are in the field, then $K - K_0 = K_1$ is a homogeneous formal covariant of S which is congruent to zero whenever the coefficients are in the field—that is, K_1 vanishes for all classes of forms of the system S. There are two possibilities: (1) K_1 is the product of two or more rational formal covariants of which none vanishes for all sets of coefficients in the field; (2) K_1 contains as a factor a formal covariant which can not be expressed as the product of several covariants and which does vanish for all classes of S.

Thus any formal covariant K of the system S is expressible in one of the two following forms:

$$K = K_0 + M_1 M_2 \cdots M_r, \tag{1}$$

where $M_1, M_2, \cdots, M_r$ are formal covariants of S which do not vanish for all sets of values of the coefficients in the field;

$$K = K_0 + VK_1 \tag{2}$$

in which V and K_1 are formal covariants of S where V vanishes whenever the $a, b, c, \cdots$ are in the field and is not the product of two or more formal covariants of S. Notice, moreover, that we can use the same formal covariant K_0 for all formal covariants which are of the same degree in $a, b, c, \cdots$ and of the same order in x and y and which, moreover, are congruent to the same modular covariant C when $a, b, c, \cdots$ are in the field.

Now let Σ denote a set of formal covariants of S determined in the following manner. Consider any particular modular covariant C of S (in which all the exponents of $a, b, c, \cdots$ are $\leq p^n - 1$) and the totality of all formal covariants K which are congruent to C whenever the coefficients are marks of the field. The degrees of these covariants K will be of the form $\omega + q(p^n - 1)$ where q is a positive integer and ω is the least such degree. For each such degree, choose one formal covariant K to put in the set Σ.* Do this for every

* It must be borne in mind that there are not necessarily any formal covariants for *each* such degree. Compare the third footnote in § 2.

modular covariant C of a fundamental set of modular covariants of S, and let Σ denote the set of all such formal covariants.

Then, proceeding by induction, we see that every formal covariant of S is a polynomial in the covariants of the set Σ and in those irreducible formal covariants of S which vanish whenever the coefficients $a, b, c, \cdots$ are marks of the field. Thus we have proved

THEOREM III. *Every formal modular covariant of a system S of binary forms with respect to the Galois Field $GF[p^n]$ is a polynomial in the modular covariants which have been made formally invariant as to the coefficients of the forms, and in the irreducible covariants which are congruent to zero whenever the coefficients are marks of the field.*

It will now be interesting to give a more elegant proof of this theorem by the aid of a symbolic notation.

6. Symbolic. Notation. In the symbolic theory of algebraic invariants, Aronhold, Clebsch and Gordan* express any binary form f of order m as the mth power of a symbolic linear function of the variables thus,—

$$f = a_0 x^m + a_1 x_1^{m-1} x_2 + \cdots = (\alpha_1 x_1 + \alpha_2 x_2)^m = \alpha_x^m .$$

Then $a_0 = \alpha_1^m$, $a_1 = m\alpha_1^{m-1} \alpha_2$, and in general $a_r = {}_mC_r \alpha_1^{m-r} \alpha_2^r$. In order that the a's may be independent variables, the agreement is made that we never use any term of more than the mth degree in the α's. Accordingly, if we wish to express a product of two or more a's in symbolic form, we have to introduce two or more equivalent sets of symbols, say (α_1, α_2), (β_1, β_2), etc. Then $f = \alpha_x^m = \beta_x^m = \gamma_x^m = \cdots$; and thus we can write $a_0 a_2$ as $\alpha_1^m (m\beta_1^{m-1} \beta_2)$ or as $\beta_1^m (m\alpha_1^{m-1} \alpha_2)$, etc. Then every polynomial in these symbols which has the invariantive property will be an invariant of f; though, in order that such an invariant may be rational and integral in the a's it must be homogeneous of the mth degree in the α's, homogeneous of the mth degree in the β's, etc.

In the theory of modular invariants (formal and otherwise), we can not, however, use this classical symbolic notation for the binary form f over the general Galois field $GF[p^n]$. In the first place, the binomial coefficients which naturally arise in this way may be zero in the field. In the second place, we know that $L_a = (\alpha^{p^n} \alpha) = \alpha_1^{p^n} \alpha_2 - \alpha_1 \alpha_2^{p^n}$ is an invariant symbol, since $\alpha_1^{p^n}, \alpha_2^{p^n}$ are cogredient with α_1, α_2. But in case $p^n + 1 > m$, we have no right to use this symbol, as explained above. Accordingly we must adopt some other symbolism.

[*] Aronhold, Journal für Mathematik, vol. 62 (1863), p. 281; Clebsch, Journal für Mathematik, vol. 59 (1860), p. 1; Gordan, Mathematische Annalen, vol. 2 (1870), p. 227, vol. 5 (1872), p. 595; Clebsch, *Theorie der binären algebraischen Formen* (1872); Gordan, *Vorlesungen über Invariantentheorie* (1887).

With this in mind, represent the binary form as

$$f = \prod(\alpha_1 x_1 + \alpha_2 x_2) = (\alpha_1 x_1 + \alpha_2 x_2)(\beta_1 x_1 + \beta_2 x_2) \cdots = \prod \alpha_x \beta_x \cdots,$$

where it is understood that there are m distinct symbolic factors. Now every invariantive function of the variables and the coefficients of f is an invariantive function of the symbols $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \cdots$ and x_1, x_2. Conversely, an invariantive function C of the symbols is a symbolic covariant of f; and if it is rational in the a's it is actually a covariant. If, however, C is to be rational in the a's it must be such that, if we interchange any two pairs of symbols—say (α_1, α_2) and (β_1, β_2)—it is unchanged in form, and every term of C must be of the same degree in the α's that it is in the β's, etc. If these two conditions are satisfied, then conversely C is rational in the a's. For C is then a symmetric function of the pairs $(\alpha_1, \alpha_2), (\beta_1, \beta_2), \cdots$ and hence, by the theory of symmetric functions, is well-known to be rational in a certain finite set of such functions, which are simply the a's.

7. Symbolic proof of the theorem of section 5. As in the theory of algebraic invariants, we notice that $(\alpha_1, \alpha_2), (\beta_1, \beta_2), \cdots$ are what we might call pseudo-cogredient with $(x_2, -x_1)$ that is, $(\alpha_1, \alpha_2), (\beta_1, \beta_2), \cdots$ are cogredient with $(x_2, -x_1)$ aside from a power of the modulus. For, if x_1 and x_2 are subjected to the non-singular transformation

$$\begin{cases} x_1 = ax_1' + bx_2', \\ x_2 = cx_1' + dx_2', \end{cases} \qquad \Delta = ad - bc \neq 0,$$

then, since $\alpha_1 x_1 + \alpha_2 x_2 = \alpha_1' x_1' + \alpha_2' x_2'$,

$$\begin{cases} \alpha_2 = \dfrac{1}{\Delta}[a\alpha_2' + b(-\alpha_1')], \\ -\alpha_1 = \dfrac{1}{\Delta}[c\alpha_2' + d(-\alpha_1')]. \end{cases}$$

Hence every modular covariant of a system S of forms is a modular invariant of certain pairs $(\alpha_1, \alpha_2), (\beta_1, \beta_2), \cdots, (x_2, -x_1)$ which are pseudo-cogredient; or every modular covariant of the system S may be regarded as a modular invariant of certain cogredient pairs $(\alpha_1, \alpha_2), (\beta_1, \beta_2), \cdots, (x_2, -x_1)$. The converse is also true, though not every invariant of the symbolic pairs is a rational covariant of the system S. As proved at the end of § 6, a necessary and sufficient condition that an invariant of the symbolic pairs be rational is that it be symmetric in these pairs. A similar remark applies to the covariants.

In fact, by the proof of the fundamental theorem, any formal modular covariant C of a single binary form f is of the form $M + V$, where M is a modular invariant of the $n + 1$ pairs which has been made formally invariant and V is a formal invariant of the $n+1$ pairs which vanishes whenever the α's, β's, $\cdots$ are all in the field. This M can be taken as the same for all covariants C which are congruent to the same modular covariant when the α's, β's, $\cdots$ are all in the field.

It can readily be proved that we can so choose M and V that they are symmetric in the pairs (α_1, α_2), (β_1, β_2), $\cdots$. For, by the proof of the fundamental theorem, $C = M_\alpha + L_\alpha M_{1\alpha}$, where M_α and $M_{1\alpha}$ are two formal modular invariants of the $n+1$ pairs. Similarly $C = M_\beta + L_\beta M_{1\beta}$, etc. Now M_β and $M_{1\beta}$ can be so chosen that they are obtained from M_α and $M_{1\alpha}$ respectively by interchanging the pairs (α_1, α_2) and (β_1, β_2). Hence

$$C = M_{\alpha\beta} + (L_\alpha N_\alpha + L_\beta N_\beta + L_\alpha L_\beta N_{\alpha\beta}),$$

where N_β is obtained from N_α by interchanging the pairs (α_1, α_2) and (β_1, β_2) and where $M_{\alpha\beta}$ and $N_{\alpha\beta}$ are symmetric in these two pairs. By induction, we see that the statement at the beginning of this paragraph is true.

With slight changes, this proof holds for a system of binary forms.

It is to be noted that the symbol L_α vanishes whenever α_1 and α_2 are marks of the field; similarly with L_β, etc. But any formal modular invariant which evanesces whenever the α's, β's, etc. are all in the field (such as those of the type $L_\alpha N_\alpha + L_\beta N_\beta + L_\alpha L_\beta N_{\alpha\beta} + \cdots$) does not necessarily vanish whenever the coefficients of the system S are in the field. Hence we have proved Theorem III.

This second proof, besides beauty of form, has the advantage of indicating the relation between formal modular covariants of two forms of different degrees. Although it has the obvious disadvantage that it does not furnish a definite formula by which we can determine the formal modular covariants of a system of forms, nevertheless it suggests that some day there may be evolved a symbolic theory of formal covariants.

8. Application to the binary quadratic, modulo 3. Dickson[*] has shown that a fundamental set of modular invariants of the binary quadratic,

$$f_2 = a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2,$$

modulo 3 is

$$\Delta = a_1^2 - a_0 a_2, \qquad q = (a_0 + a_2)(a_1^2 + a_0 a_2 - 1);$$

while he has shown[†] that a fundamental set of formal invariants consists of

$$\Delta = a_1^2 - a_0 a_2, \qquad J = a_0(a_0 + a_1 + a_2)(a_0 + 2a_1 + a_2)a_2,$$

$$B = a_1(a_1 + a_0)(a_1 - a_0)(2a_0 + a_2)(2a_1 + a_2)(a_1 + a_2),$$

$$\Gamma = (a_0 + a_2)(2a_0 + 2a_1 + a_2)(2a_0 + a_1 + a_2).$$

[*] These Transactions, vol. 8 (1907), p. 209.

[†] These Transactions, vol. 14 (1913), p. 310.

To verify Theorem III for the invariants of f_2, we compute the values of q, Δ, J, B and Γ for the different classes and find that

Class	f_2	q	Δ	J	B	Γ
1	0	0	0	0	0	0
2	x^2	2	0	0	0	1
3	$-x^2$	1	0	0	0	2
4	x^2+y^2	0	2	1	0	0
5	$2xy$	0	1	0	0	0

Hence, whenever the a's are in the field, $\Gamma \equiv -q$, $J \equiv 2\Delta^2 + \Delta$, $B \equiv 0$.

Dickson† has also shown that a fundamental set of modular covariants consists of

$$q,\ \Delta,\ L,\ Q,\ f_2,\ f_4 = a_0 x_1^4 + a_1 (x_1^3 x_2 + x_1 x_2^3) + a_2 x_2^4,$$

$$C_1 = (a_0^2 a_1 - a_1^3) x_1^2 + (a_0 - a_2)(a_1^2 + a_0 a_2) x_1 x_2 + (a_1^3 - a_1 a_2^2) x_2^2,$$

$$C_2 = (a_0^2 + \Delta) x_1^2 + a_1 (a_0 + a_2) x_1 x_2 + (a_2^2 + \Delta) x_2^2.$$

Glenn* later showed that a fundamental set of formal covariants consists of eighteen forms:

$$\Delta,\ J,\ B,\ \Gamma,\ L,\ Q,\ f_2,\ f_4,\ f_6 = a_0 x_1^6 + 2a_1 x_1^3 x_2^3 + a_2 x_2^6,$$

$$C_1,\ C_2,\ C_4 = (a_0^2 + \Delta) x_1^4 + 2a_1 (a_0 + a_2)(x_1^3 x_2 + x_1 x_2^3) + (a_2^2 + \Delta) x_2^4,$$

$$\zeta_4 = (a_0^2 a_1 - a_1^3) x_1^4 - (a_0 - a_2)(a_1^2 + a_0 a_2)(x_1^3 x_2 + x_1 x_2^3) + (a_1^3 - a_1 a_2^2) x_2^4,$$

$$-\zeta_6 = (a_0^2 a_1 - a_1^3) x_1^6 + (a_0 - a_2)(a_1^2 + a_0 a_2) x_1^3 x_2^3 + (a_1^3 - a_1 a_2^2) x_2^6,$$

$$\phi_2 = a_0^3 x_1^2 + 2a_1^3 x_1 x_2 + a_2^3 x_2^2,$$

$$\phi_4 = a_0^3 x_1^4 + a_1^3 (x_1^3 x_2 + x_1 x_2^3) + a_2^3 x_2^4,$$

$$\vartheta_2 = (a_0 a_1^3 - a_0^3 a_1) x_1^2 + (a_0 a_2^3 - a_0^3 a_2) x_1 x_2 + (a_1 a_2^3 - a_1^3 a_2) x_2^2,$$

$$\xi_2 = (a_0^2 a_1^3 - a_0^4 a_1) x_1^2 + (a_0^3 a_1^2 + 2a_0^4 a_2 + 2a_0 a_1^4 + a_1^4 a_2 + 2a_0^3 a_2^2 + a_0^2 a_2^3 + 2a_1^2 a_2^3 + a_0 a_2^4) x_1 x_2 + (a_1 a_2^4 - a_1^3 a_2^2) x_2^2.$$

Of the formal covariants not in the fundamental set of modular covariants, $f_6 \equiv f_2^3$, $C_4 \equiv (\Delta + 1) f_2^2 + C_1^2$, $-\zeta_6 \equiv C_1^3$, $\zeta_4 \equiv C_1 C_2$, $\phi_2 \equiv f_2$ and $\phi_4 \equiv f_4$ whenever the a's are in the field. Finally, $\vartheta_2 \equiv \xi_2 \equiv 0$ whenever the a's are in the field. Thus the theorem is verified.

* These Transactions, vol. 20 (1919), pp. 154–168. For convenience, we have used C_4 instead of D_4, which is legitimate, since $D_4 = C_4 - f_2^2$.

The writer has also applied the theorem to the set of fundamental formal covariants of the binary cubic, modulo 2, which has been found by Glenn.* We leave the details to the reader.

9. A general application. Theorem III enables us at once to narrow down the question of the finiteness of formal covariants of a system S of forms with respect to the Galois field $GF[p^n]$. There is no loss of generality to consider only homogeneous formal covariants and invariants.

Let Q_1 be a homogeneous formal invariant of lowest degree q which is $\equiv 1$ whenever the coefficients are in the field, but not all zero. Now there are polynomials in the coefficients which are $\equiv 1$ for all sets of values of coefficients in the field not all zero and which are formally invariant. For let $I_1, I_2, \cdots, I_\nu$ be the characteristic modular invariants for all classes except the one in which all forms are identically zero, and let $I'_1, I'_2, \cdots, I'_\nu$ be formal invariants of lowest degree which are congruent to $I_1, I_2, \cdots, I_\nu$ respectively whenever the coefficients are in the field. Let the degrees of $I'_1, I'_2, \cdots, I'_\nu$ be $d_1, d_2, \cdots, d_\nu$ respectively. Then, if D be the least common multiple of the d's,

$$Q' = I_1'^{\delta_1} + I_2'^{\delta_2} + \cdots + I_\nu'^{\delta_\nu} \qquad (\delta_i = D/d_i)$$

is a homogeneous formal invariant which is congruent to 1 whenever the coefficients are marks of the field not all zero.

Now let C be any modular covariant of the system S of order ω, and let C_1 be a formal covariant which is congruent to C whenever the coefficients are in the field; moreover, let C_1 be such a covariant of order ω and of lowest degree in the coefficients. Also, let $C_1, C_2, \cdots, C_k$ be formal covariants of order ω which are congruent to C for all sets of the coefficients in the field and which are of all possible degrees ranging from c_1 up to but not including $c_1 + q$. Let their degrees be respectively $c_1, c_2, \cdots, c_k$.

If K_i be any formal covariant of order ω which is congruent to C whenever the coefficients are in the field, but not all zero and which is of degree c_i, then K_i is identically equal to $C_i +$ (a formal covariant which is congruent to zero whenever the coefficients are in the field).

Now $C_1 Q_1, \cdots, C_k Q_1$ are all formal covariants which are congruent to C for all sets of coefficients in the field. The degrees of $C_1 Q_1, \cdots, C_k Q_1$ are $c_1 + q, c_2 + q, \cdots, c_k + q$ and range from $c_1 + q$ up to but not including $c_1 + 2q$. If there be any formal covariants of order ω which are congruent to C and which are of a degree different from $c_1 + q, c_2 + q, \cdots, c_k + q$ and yet whose degree lies between $c_1 + q$ and $c_1 + 2q$, select one representative covariant of each such possible degree. Let these additional covariants be denoted by $A_1, \cdots, A_l$. Since there is only a finite number of integers

* These Transactions, vol. 19 (1918), pp. 109-118. The invariants had already been found by Dickson (*Madison Colloquium Lectures*, p. 56).

between $c_1 + q$ and $c_1 + 2q$, there is only a finite number of such additional representative covariants.

Thus if K' be any formal covariant of order ω which is congruent to C whenever the coefficients are in the field and which is of a degree between $c_1 + q$ and $c_1 + 2q$, then K' is identically equal to (some C_i) $\times Q_1 +$ (a formal covariant which is congruent to zero whenever the coefficients are in the field); or K' is identically equal to (some A_i) $+$ (a formal covariant which is congruent to zero whenever the coefficients are in the field).

Proceed similarly with the formal covariants of order ω which are congruent to C whenever the coefficients are in the field and which are of a degree between $c_1 + 2q$ and $c_1 + 3q$; between $c_1 + 3q$ and $c_1 + 4q$; etc.

Thus we have a set of those formal covariants of order ω which are congruent to C whenever the coefficients are in the field which is such that there is one and only one such covariant for each such possible degree. Moreover, in view of the preceding argument and in view of the fact that there is only a finite number of additional representative covariants (such as $A_1, \cdots, A_l$, etc.), all the representative covariants of order ω which are congruent to C are expressible as the product of a power of Q_1 by one of a finite number of these representative covariants.

Combining these results, we see that all formal covariants of order ω which are congruent to C for all sets of the coefficients in the field are expressible in terms of Q_1, a finite number of such covariants, and the formal covariants of order ω which are congruent to zero for all sets of the coefficients in the field.

By a similar line of reasoning, we can deduce that all formal covariants of degree d which are congruent to C whenever the coefficients are in the field are expressible in the form (one of a finite subset of such covariants) $\times$ (a power of Q) $+$ (a formal covariant which is congruent to zero whenever the coefficients are in the field).

If we now combine these two results, and make use of the fact* that the set of all modular covariants possesses the finiteness property, we then prove by induction

THEOREM IV. *The set of all formal covariants of a system S of binary forms with respect to the Galois field $GF[p^n]$ is of such a nature that every such covariant is expressible as a polynomial in Q, Q_1, members of a finite subset of formal covariants of S and the irreducible covariants which are congruent to zero whenever the coefficients are in the field.*

Thus we reach the conclusion that the set of all formal covariants of a system S possesses the finiteness property if and only if the set of all irreducible covariants which are congruent to zero for all sets of coefficients in the field possesses the finiteness property.

MOUNT HOLYOKE COLLEGE,
SOUTH HADLEY, MASS.

* See section 2 of this paper.

ON THE CONVERGENCE OF CERTAIN TRIGONOMETRIC AND POLYNOMIAL APPROXIMATIONS[*]

BY

DUNHAM JACKSON

1. **Introduction.** Let $f(x)$ be a continuous function of period 2π. Let $T_{mn}(x)$ be the trigonometric sum of order n or lower for which the integral

$$\int_0^{2\pi} |f(x) - T_{mn}(x)|^m \, dx$$

has the smallest possible value, when m and n are given. The writer has recently proved the existence and uniqueness of $T_{mn}(x)$ for any positive integral value of n and for any value of $m \geqq 1$, whether integral or not,[†] and has given some indication of the behavior of $T_{mn}(x)$, for fixed n, as m becomes infinite.[‡]

The first purpose of the present paper is to discuss the convergence of $T_{mn}(x)$ toward the value $f(x)$, when m is held fast and n is allowed to become infinite. It is found to be a sufficient condition for uniform convergence that $\lim_{\delta=0} \omega(\delta)/\sqrt[m]{\delta} = 0$, where $\omega(\delta)$ is the modulus of continuity[§] of the function $f(x)$, that is, the maximum of $|f(x') - f(x'')|$ for $|x' - x''| \leqq \delta$. The proof makes use of Bernstein's theorem on the derivative of a trigonometric sum. For $m = 1$, the condition as stated would require that $f(x)$ be a constant, but in this case it is sufficient that $f(x)$ have a continuous derivative.

One would be inclined to expect that convergence would be more likely for large than for small values of m, and the result obtained is favorable to this view, as far as it goes. The theorem is less general, however, even for large

[*] Presented to the Society, September 8 and October 30, 1920.

[†] For $m > 1$, see D. Jackson, *On functions of closest approximation*, these Transactions, vol. 22 (1921), pp. 117–128; for $m = 1$, D. Jackson, *Note on a class of polynomials of approximation*, submitted to these Transactions. These papers will be referred to by the letters A and B respectively. In A, the treatment applies to a general class of approximating functions, including trigonometric sums as a very special case; in B, the discussion relates primarily to polynomials, but attention is called to the fact that the case of trigonometric sums is to be dealt with in the same way.

[‡] Loc. cit. A, §§ 8, 9.

[§] For the name, cf. de la Vallée Poussin, *Leçons sur l'approximation des fonctions d'une variable réelle*, Paris, 1919, pp. 7–8.

values of m, than the well-known Lipschitz-Dini condition in the case of Fourier's series, $m = 2$, and the present results may for this reason be regarded as only preliminary.

The latter part of the paper contains a discussion of the corresponding problem in polynomial approximation. The method of treatment is generally the same, though the details of the reasoning and the conclusions reached are not exactly parallel in the two cases.

To return to the trigonometric case, it will be supposed throughout the discussion that m has a fixed value $\geqq 1$, given in advance. For the sake of simplicity, the subscript m will be omitted, and $T_{mn}(x)$ will be denoted simply by $T_n(x)$. This particular trigonometric sum will be referred to briefly as the *approximating function of order* n.

2. **Reduction of the problem.** If $T_n(x)$ is the approximating function for $f(x)$, and $t_n(x)$ is an arbitrary trigonometric sum of order n or lower, the approximating function for $f(x) + t_n(x)$ will be $T_n(x) + t_n(x)$. For the error is identically the same in both cases, and the integral of the mth power of its absolute value is the same, and if it were possible to diminish the value of the integral in one case, it could be diminished in the other case as well.*

Suppose there exists a trigonometric sum $t_n(x)$ such that

$$|f(x) - t_n(x)| \leqq \epsilon$$

for all values of x, where ϵ is some small positive quantity. Let

$$\phi(x) = f(x) - t_n(x).$$

Then the error of the approximating function for $\phi(x)$ is the same as the error of the approximating function for $f(x)$, and the problem of determining the magnitude of the error in the case of f is reduced to the corresponding problem for ϕ.

3. **Relation between maximum and integral.** Let ϵ be an arbitrary positive quantity; let $\phi(x)$ be an arbitrary continuous function of period 2π, subject to the condition that

$$|\phi(x)| \leqq \epsilon$$

for all values of x; and let $t(x)$ be an arbitrary trigonometric sum of order not exceeding n. Throughout this section, n will be fixed in value, and need not be indicated as a subscript. Let μ be the maximum of $|t(x)|$, and let x_0 be a point such that $|t(x_0)| = \mu$. Finally, let it be supposed that $\mu \geqq 4\epsilon$. If $\mu < 4\epsilon$, this fact in itself will serve the purpose for which the conclusion of the present section is to be used in the contrary case.

* The relation between given function and approximating function is of course not linear in general; that is, the approximating function for the sum of two given functions is not generally the sum of the approximating functions.

A theorem of S. Bernstein,* the importance of which is only beginning to be realized, states that the maximum of the absolute value of the derivative of a trigonometric sum of order n can not exceed n times the maximum of the absolute value of the original sum itself. That is, in the present case,

$$|t'(x)| \leqq \mu n,$$

for all values of x. When

$$|x - x_0| \leqq \frac{1}{2n}, \tag{1}$$

therefore,

$$|t(x) - t(x_0)| \leqq \tfrac{1}{2}\mu,$$

by the mean value theorem, and

$$|t(x)| \geqq \tfrac{1}{2}\mu.$$

Since, furthermore, $|\phi(x)| \leqq \epsilon$ and $\epsilon \leqq \mu/4$,

$$|t(x) - \phi(x)| \geqq \tfrac{1}{4}\mu$$

throughout the interval (1), of length $1/n$. From this it follows that

$$\int_0^{2\pi} |t(x) - \phi(x)|^m \, dx \geqq \frac{1}{n} \cdot \left(\frac{\mu}{4}\right)^m.$$

4. **Proof of convergence.** Now let $\tau_n(x)$ be the approximating function of order n for $\phi(x)$, let μ_n be the maximum of $|\tau_n(x)|$, and let

$$\gamma_n = \int_0^{2\pi} |\tau_n(x) - \phi(x)|^m \, dx.$$

If $\tau_n(x)$ were replaced by zero in the last integral, the value of the integral would be less than or equal to $2\pi\epsilon^m$, since $|\phi| \leqq \epsilon$. But $\tau_n(x)$ has the property of giving the integral its minimum value, and so it is certain that

$$\gamma_n \leqq 2\pi\epsilon^m.$$

By the preceding section, on the other hand, if $\mu_n \geqq 4\epsilon$,

$$\gamma_n \geqq \frac{1}{n} \cdot \left(\frac{\mu_n}{4}\right)^m.$$

So

$$\frac{1}{n} \cdot \left(\frac{\mu_n}{4}\right)^m \leqq 2\pi\epsilon^m,$$

$$\mu_n \leqq 4\epsilon(2\pi)^{1/m} n^{1/m}.$$

* See, e.g., de la Vallée Poussin, op. cit., pp. 39–42; also S. Bernstein, *Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné*, Mémoire couronné, Brussels, 1912, p. 20; and de la Vallée Poussin, Comptes Rendus, vol. 166 (1918), pp. 843–846. In Bernstein's original statement of the theorem, the factor introduced on differentiating was $2n$, instead of n.

But if $\mu_n < 4\epsilon$, the last inequality is satisfied *a fortiori*, and the relation is therefore true generally. To summarize:

If $\phi(x)$ is a continuous function of period 2π which never exceeds ϵ in absolute value, $\tau_n(x)$ the approximating function of order n for $\phi(x)$, and μ_n the maximum of $|\tau_n(x)|$, then

$$\mu_n \leqq k_m\, n^{1/m}\, \epsilon,$$

where k_m is a constant depending only on m.

Let ρ_n be the maximum of $|\tau_n(x) - \phi(x)|$; then ρ_n satisfies an inequalıty of the same form as μ_n. For $\rho_n \leqq \mu_n + \epsilon$, and so

$$\rho_n \leqq (k_m + 1)\, n^{1/m}\, \epsilon.$$

In consequence of § 2, the last assertion has the following further significance:

If $f(x)$ is a continuous function of period 2π which can be represented by a trigonometric sum of order n or lower with an error not exceeding ϵ_n, if $T_n(x)$ is the approximating function of order n for $f(x)$, corresponding to the exponent m, and if ρ_n is the maximum of $|f(x) - T_n(x)|$, then

$$\rho_n \leqq (k_m + 1)\, n^{1/m}\, \epsilon_n, \tag{2}$$

where k_m has the same meaning as before.

Hence $T_n(x)$ will converge uniformly to the value $f(x)$, if $f(x)$ can be represented by trigonometric sums for successive values of n in such a way that

$$\lim_{n=\infty} \sqrt[m]{n} \cdot \epsilon_n = 0.$$

Let $\omega(\delta)$ be the maximum of $|f(x') - f(x'')|$ for $|x' - x''| \leqq \delta$. Then there will exist representations* of $f(x)$ such that $\epsilon_n \leqq c\omega(2\pi/n)$, where c is independent of n, for all positive integral values of n. *If*

$$\lim_{\delta=0} \omega(\delta)/\sqrt[m]{\delta} = 0,$$

then $\sqrt[m]{n/2\pi} \cdot \omega(2\pi/n)$ will approach zero as n becomes infinite, and the condition for convergence will be satisfied.

In the case $m = 1$, the condition $\lim_{\delta=0} \omega(\delta)/\delta = 0$ would require that $f(x)$ be a constant. *There will exist representations† of $f(x)$ such that*

$$\lim_{n=\infty} n\epsilon_n = 0,$$

however, provided that $f(x)$ has a continuous derivative, so that the last-named condition is sufficient for convergence in this case.

If $\omega(\delta)$ approaches zero more rapidly than is required by the criterion for

* Cf., e.g., D. Jackson, *On the approximate representation of an indefinite integral*, etc., these Transactions, vol. 14 (1913), pp. 343–364; p. 350.

† Cf., e.g., the paper cited in the preceding footnote, pp. 350–351.

convergence, the relation (2), together with such general theorems on trigonometric approximation as are contained in the paper just cited, will give information with regard to the degree of convergence of the present approximating functions.

It may be remarked that the reasoning above, as applied to the case $m = 2$, gives a new proof of the convergence of the Fourier's series for an extensive class of functions, although, in view of its dependence on Bernstein's theorem and on the general theory of approximation by trigonometric sums, the proof is not more elementary than others that are well known.

5. **Polynomial case, end-points included.** The theorem on the derivative of a trigonometric sum, on which the preceding proof is based, has a counterpart relating to polynomials. Let $P_n(x)$ be a polynomial of the nth degree or lower, such that $|P_n(x)| \leqq \mu$ for $-1 \leqq x \leqq 1$, where μ is a constant. Let $x = \cos\theta$; then $P_n(x)$ is a trigonometric sum in θ, of order n or lower, the absolute value of which is $\leqq \mu$ for all values of θ. By the theorem for the trigonometric case,

$$\left|\frac{d}{d\theta} P_n(\cos\theta)\right| = |\sin\theta P_n'(\cos\theta)| \leqq \mu n .$$

That is,*

$$|\sqrt{1-x^2}\, P_n'(x)| \leqq \mu n$$

for $-1 \leqq x \leqq 1$.

By means of a linear transformation on x, it can be inferred further that if $|P_n(x)| \leqq \mu$ for $a \leqq x \leqq b$, and if $a < a' < b' < b$, then

$$|P_n'(x)| \leqq C\mu n \tag{3}$$

for $a' \leqq x \leqq b'$, where C depends only on a, b, a', and b'.

The immediate application of this result to the problem of convergence is not apparent. It is possible to reason further, however, as follows. Let it be supposed once more that the interval in question is $(-1, 1)$. The function $P_n'(\cos\theta)$, being a polynomial of degree $n-1$ at most in $\cos\theta$, is a cosine-sum of order $n-1$ at most in θ. Since

$$\cos k\theta \sin\theta = \tfrac{1}{2}[\sin(k+1)\theta - \sin(k-1)\theta]$$

for all values of k, the expression

$$Q_n(\theta) = \sin\theta P_n'(\cos\theta)$$

has the form

$$Q_n(\theta) = A_1 \sin\theta + A_2 \sin 2\theta + \cdots + A_n \sin n\theta .$$

* S. Bernstein, loc. cit., pp. 6–11, proves this theorem directly, and makes it the basis for the discussion of the trigonometric theorem; but it is simpler to begin with de la Vallée Poussin's proof for the trigonometric case.

Let the maximum of $|Q_n(\theta)|$ be denoted by μ_1; the theorem of Bernstein for trigonometric sums is applicable again, with the conclusion that

$$|Q'_n(\theta)| \leqq \mu_1 n$$

for all values of θ. As $Q_n(0) = 0$, it follows that

$$|Q_n(\theta)| \leqq \mu_1 n |\theta|.$$

Let θ have a value in the interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$; then

$$\left|\frac{\sin\theta}{\theta}\right| \geqq \frac{2}{\pi} > \frac{1}{2}, \qquad |\theta| \leqq 2|\sin\theta|,$$

and

$$\left|\frac{Q_n(\theta)}{\sin\theta}\right| \leqq 2\mu_1 n. \tag{4}$$

The same result can be obtained for $\pi/2 \leqq x \leqq 3\pi/2$ from the relations

$$Q_n(\pi) = 0, \qquad |Q_n(\theta)| \leqq \mu_1 n|\theta - \pi|,$$
$$|\theta - \pi| \leqq 2|\sin\theta|.$$

The relation (4) therefore holds for all values of θ. But

$$\left|\frac{Q_n(\theta)}{\sin\theta}\right| = P'_n(\cos\theta) = P'_n(x),$$

and, on the other hand, $\mu_1 \leqq \mu n$. Hence it follows that*

$$|P'_n(x)| \leqq 2\mu n^2$$

for $-1 \leqq x \leqq 1$. If $|P_n(x)| \leqq \mu$ for $a \leqq x \leqq b$, the conclusion is that

$$|P'_n(x)| \leqq \frac{4\mu n^2}{b-a} \tag{5}$$

throughout the same interval.

Now let $f(x)$ be a given continuous function for $a \leqq x \leqq b$, let m be a given fixed number $\geqq 1$, and let $P_n(x)$ be the polynomial of the nth degree or lower for which

$$\int_a^b |f(x) - P_n(x)|^m \, dx$$

is a minimum. For reasons analogous to those which appeared in connection with the trigonometric problem, if $p_n(x)$ is a polynomial of degree n or lower, such that

$$|f(x) - p_n(x)| \leqq \epsilon$$

* This result is contained in a theorem of A. Markoff (see S. Bernstein, loc. cit., pp. 11–13), which states that $|P_n(x)| \leqq \mu n^2$ throughout the interval. If it could be shown above that $|Q_n(\theta)/(\sin\theta)| \leqq \mu_1 n$, the present method would prove Markoff's theorem completely.

throughout the interval, ϵ being small, the error of the approximating polynomial $P_n(x)$ for $f(x)$ is the same as the error of the approximating polynomial for

$$\phi(x) = f(x) - p_n(x),$$

and the latter is easier to deal with.

Suppose, then, that $\phi(x)$ is an arbitrary continuous function such that

$$|\phi(x)| \leqq \epsilon$$

throughout (a, b). Let $\pi_n(x)$ be the approximating polynomial of the nth degree for $\phi(x)$, let $\mu_n = |\pi_n(x_0)|$ be the maximum of $|\pi_n(x)|$ in (a, b), and let it be assumed for the moment that $\mu_n \geqq 4\epsilon$.

In consequence of (5), if

$$|x - x_0| \leqq \frac{b-a}{8n^2},$$

it will certainly be true that

$$|\pi_n(x) - \pi_n(x_0)| \leqq \tfrac{1}{2}\mu_n,$$

$$|\pi_n(x)| \geqq \tfrac{1}{2}\mu_n,$$

$$|\pi_n(x) - \phi(x)| \geqq \tfrac{1}{4}\mu_n.$$

The last relation holds throughout an interval of length at least $(b-a)/(8n^2)$, and it is therefore certain that

$$\gamma_n = \int_a^b |\pi_n(x) - \phi(x)|^m\,dx \geqq \frac{b-a}{8n^2}\cdot\left(\frac{\mu_n}{4}\right)^m.$$

On the other hand, γ_n is at any rate not greater than the value that would be obtained if $\pi_n(x)$ in the integral were replaced by zero, that is,

$$\gamma_n \leqq (b-a)\epsilon^m.$$

Hence

$$\frac{(b-a)}{8n^2}\cdot\left(\frac{\mu_n}{4}\right)^m \leqq (b-a)\epsilon^m,$$

$$\mu_n \leqq 4\epsilon\cdot 8^{1/m}n^{2/m}. \tag{6}$$

The last relation has been obtained on the hypothesis that $\mu_n \geqq 4\epsilon$, but it is clearly satisfied also if $\mu_n < 4\epsilon$.

Let ρ_n be the maximum of $|\pi_n(x) - \phi(x)|$; then $\rho_n \leqq \mu_n + \epsilon$, while, at the same time, if $\phi(x)$ has the form $f(x) - p_n(x)$, ρ_n is the maximum of the difference between $f(x)$ and its approximating polynomial $P_n(x)$. Thus the following assertion is justified:

If $f(x)$ is continuous for $a \leqq x \leqq b$, and can be represented throughout the interval by a polynomial of degree n or lower with an error not exceeding ϵ_n, if

$P_n(x)$ is the approximating polynomial of degree n for $f(x)$, corresponding to the exponent m, and if ρ_n is the maximum of $|f(x) - P_n(x)|$, then

$$\rho_n \leqq c_m n^{2/m} \epsilon_n,$$

where c_m is a constant depending only on m.

The constant c_m corresponds to that which was denoted by $k_m + 1$ in the trigonometric case; the difference in notation has no significance.

The polynomial $P_n(x)$ will converge uniformly to the value $f(x)$ for $a \leqq x \leqq b$, if it is possible to make ϵ_n take on such values that

$$\lim_{n=\infty} \sqrt[m]{n^2} \cdot \epsilon_n = 0.$$

Let $\omega(\delta)$ be the modulus of continuity of $f(x)$, as before. There will exist polynomial representations* of $f(x)$ such that $\epsilon_n \leqq c' \omega[(b-a)/n]$, where c' is independent of n, for all positive integral values of n. *A sufficient condition for uniform convergence throughout the interval $a \leqq x \leqq b$ is that $\lim_{\delta=0} \omega(\delta)/\sqrt[m]{\delta^2} = 0$.*

For $m \leqq 2$, the condition just stated would require that $f(x)$ be a constant, but it is easy to replace this condition by a more satisfactory one. If $1 < m \leqq 2$, let it be assumed that $f(x)$ possesses a continuous first derivative having the modulus of continuity $\omega_1(\delta)$. It is possible to make†

$$\epsilon_n \leqq \frac{c''}{n} \omega_1\left(\frac{b-a}{n-1}\right),$$

where c'', once more, is independent of n. *For $1 < m \leqq 2$, a sufficient condition for uniform convergence throughout the interval $a \leqq x \leqq b$ is that $f(x)$ have a continuous derivative, and that $\lim_{\delta=0} \omega_1(\delta)/\delta^{(2/m)-1} = 0$, where $\omega_1(\delta)$ is the modulus of continuity of the derivative.* In case $m = 2$, the condition thus obtained is simply that $f(x)$ have a continuous first derivative.

For $m = 1$, a sufficient condition is that $f(x)$ have a continuous second derivative.‡

If $f(x)$ possesses a greater degree of regularity than is required for mere convergence, theorems on the degree of convergence of the approximating polynomials can be written down at once; it is not necessary to dwell on the point further.

6. **Polynomial case, end-points excluded.** The relation (6) once having been obtained, it is possible to make use of (3) to some advantage in connection with the convergence at interior points of the interval.

* See, e.g., D. Jackson, these Transactions, vol. 14, as already cited, pp. 353–354.

† D. Jackson, these Transactions, vol. 14, as cited, p. 354.

‡ Cf. preceding citation.

Let the notation leading up to (6) be kept in force, let a' and b' be two numbers such that

$$a < a' < b' < b,$$

and let $\mu_n' = |\pi_n(x_1)|$ be the maximum of $|\pi_n(x)|$ for $a' \leqq x \leqq b'$, with the assumption, temporarily, that $\mu_n' \geqq 4\epsilon$.

By (6) and (3),

$$|\pi_n(x) - \pi_n(x_1)| \leqq 4C\epsilon \cdot 8^{1/m} n^{(2/m)+1} |x - x_1| = C' \epsilon \cdot n^{(2/m)+1} |x - x_1|,$$

as long as x is in (a', b'); the constant C' depends only on m, a, b, a', and b'. If

$$(7) \qquad |x - x_1| \leqq \frac{\mu_n'}{2C' \epsilon \cdot n^{(2/m)+1}},$$

x will surely be in (a', b'), on one side of x_1 at least, for $n \geqq n_0(m, a, b, a', b')$, because the right-hand member of (7), in consequence of (6) and the fact that $\mu_n' \leqq \mu_n$, can not exceed $2 \cdot 8^{1/m}/(C'n)$; and it will follow further that

$$|\pi_n(x) - \pi_n(x_1)| \leqq \tfrac{1}{2}\mu_n',$$

$$|\pi_n(x)| \geqq \tfrac{1}{2}\mu_n',$$

$$|\pi_n(x) - \phi(x)| \geqq \tfrac{1}{4}\mu_n'.$$

These relations hold throughout an interval, the length of which is at least equal to the right-hand member of (7).

This means that

$$\gamma_n = \int_a^b |\pi_n(x) - \phi(x)|^m dx \geqq \frac{\mu_n'}{2C' \epsilon \cdot n^{(2/m)+1}} \cdot \left(\frac{\mu_n'}{4}\right)^m.$$

But it is still true that

$$\gamma_n \leqq (b - a)\epsilon^m.$$

Hence

$$\mu_n'^{m+1} \leqq C'' n^{(2/m)+1} \epsilon^{m+1},$$

$$\mu_n' \leqq C''' \epsilon \cdot n^{(m+2)/[m(m+1)]},$$

the coefficients C'' and C''', like the preceding C's, depending only on m, a, b, a', and b'. A relation of the same form continues to hold if $\mu_n' < 4\epsilon$.

There is perhaps no need of setting down at length the resulting theorems on convergence and degree of convergence in the interior of the original interval. The results of the section are summed up by saying that the exponent $1/m$ of the trigonometric case, which became $2/m$ in the preceding section, is replaced here by

$$\frac{1}{m} \cdot \frac{m+2}{m+1}.$$

The University of Minnesota,
Minneapolis, Minn.

DETERMINATION OF ALL GENERAL HOMOGENEOUS POLYNOMIALS EXPRESSIBLE AS DETERMINANTS WITH LINEAR ELEMENTS[*]

BY

LEONARD EUGENE DICKSON

1. The general quadratic forms in three and four variables can be transformed into $x_1 x_2 - x_3^2$ and $x_1 x_2 - x_3 x_4$ respectively, and hence are expressible as determinants of order 2. Since any binary form of degree r is a product of r linear forms, it is expressible as an r-rowed determinant whose elements outside the main diagonal are all zero.

It was proved geometrically by H. Schröter[†] and more simply by L. Cremona[‡] that a sufficiently general cubic surface $f = 0$ is the locus of the intersections of corresponding planes of three projective bundles of planes:

$$\kappa l_{i1} + \lambda l_{i2} + \mu l_{i3} = 0 \qquad (i = 1, 2, 3),$$

where κ, λ, μ are parameters and the l_{ij} are linear homogeneous functions[§] of $x_1, \cdots, x_4$. Hence $f = 0$ has the determinantal form $|l_{ij}| = 0$. Taking $x_4 = 0$, we see that a general cubic curve is expressible in determinantal form.

I shall prove that every plane curve is expressible in determinantal form and that, aside from the cases mentioned above, no further general homogeneous polynomial is expressible in determinantal form.

The case of quartic surfaces was discussed erroneously by Jessop.[||] His argument would apply equally well to the determinant D whose 16 elements are binary linear forms and show that D can be given a form containing a single parameter, whereas every binary quartic can be expressed in the form D.

A new theory of equivalence of pairs of bilinear forms is given in § 9.

2. THEOREM 1. *When the number of terms in the general form of degree r in n variables $(n > 2)$ exceeds $(n-2)r^2 + 2$, it is not expressible as a determinant whose elements are linear forms.*

[*] Presented to the Society at Chicago, December 28, 1920.

[†] Journal für Mathematik, vol. 62 (1863), p. 265.

[‡] *Ibid.*, vol. 68 (1868), p. 79.

[§] In case their coefficients are rational we obtain all rational solutions of the Diophantine equation $f = 0$ by solving our three linear equations for the ratios of $x_1, \cdots, x_4$, obtaining cubic functions of κ, λ, μ.

[||] *Quartic Surfaces*, 1916, p. 160.

Let D be any r-rowed determinant whose elements are linear homogeneous functions of $x_1, \cdots, x_n$. We may express the matrix M of D in the form $x_1 M_1 + \cdots + x_n M_n$, where each M_i is a matrix whose r^2 elements are constants. Evidently D is at most multiplied by a constant not zero if we interchange any two rows or any two columns, or multiply the elements of any row or column by a constant not zero, or add to the elements of any row (or column) the products of the elements of any other row (or column) by a constant. The effect on M of any succession of such "elementary transformations" is known to be the same as forming the product AMB, where A and B are constant matrices whose determinants are not zero.

If the determinant of M_1 is zero, D lacks x_1^r and will not represent the general form. Hence M_1 has an inverse M_1^{-1} such that $M_1 M_1^{-1}$ is the identity matrix I. Consider the new matrix

$$N = MM_1^{-1} = x_1 I + x_2 N_2 + \cdots + x_n N_n \qquad (N_i = M_i M_1^{-1}),$$

whose determinant equals the quotient of D by $|M_1|$ and has unity as the coefficient of x_1^r. The product ANB will likewise have I as the coefficient of x_1 if and only if $A = B^{-1}$. Our next step is therefore to choose matrix B so that $B^{-1} N_2 B$ shall have a canonical form. We first interpret this product. If N_2 is the matrix of the transformation

$$(1) \qquad \xi_i' = \sum_{j=1}^{r} \alpha_{ij} \xi_j \qquad (i = 1, \cdots, r),$$

and if we introduce new variables $\eta_1, \cdots, \eta_r$ by means of a transformation

$$\eta_i = \sum_{k=1}^{r} \beta_{ik} \xi_k \qquad (i = 1, \cdots, r),$$

whose matrix is B, transformation (1) becomes a transformation on the η's whose matrix is easily verified* to be $B^{-1} N_2 B$. Naturally we desire that as many as possible of the η's shall be transformed into mere multiples of themselves. Hence we seek functions $\eta = \Sigma\beta_k \xi_k$ such that $\eta' = \lambda\eta$ under transformation (1); the conditions are evidently

$$\sum_{i=1}^{r} \beta_i \alpha_{ij} = \lambda\beta_j \qquad (j = 1, \cdots, r).$$

Thus λ must be a root of the characteristic equation

$$\begin{vmatrix} \alpha_{11} - \lambda & \alpha_{12} & \cdots & \alpha_{1r} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{r1} & \alpha_{r2} & \cdots & \alpha_{rr} - \lambda \end{vmatrix} = 0$$

* Dickson, *Linear Groups*, Teubner, 1901, p. 80.

of transformation (1). When λ is any root, the above conditions are known to have solutions $\beta_1, \cdots, \beta_r$ not all zero, so that $\eta' = \lambda\eta$.

If in the last determinant we replace λ by $-x_1/x_2$ and multiply the elements of each row by x_2, we obtain

$$\begin{vmatrix} x_2\,\alpha_{11} + x_1 & x_2\,\alpha_{12} & \cdots & x_2\,\alpha_{1r} \\ \cdot & \cdot & \cdot & \cdot \\ x_2\,\alpha_{r1} & x_2\,\alpha_{r2} & \cdots & x_2\,\alpha_{rr} + x_1 \end{vmatrix},$$

which is the determinant of matrix $x_2\,N_2 + x_1\,I$ and hence is the value of D when $x_3 = 0, \cdots, x_n = 0$. The latter must be the general binary form in x_1, x_2, since D is required to represent the general form in $x_1, \cdots, x_n$. Hence the above characteristic equation must have r distinct roots $\lambda_1, \cdots, \lambda_r$.

We thus obtain r linear functions $\eta_1, \cdots, \eta_r$ which (1) multiplies by $\lambda_1, \cdots, \lambda_r$ respectively. A simple artifice* shows that these η's are linearly independent functions of $\xi_1, \cdots, \xi_r$ and hence may be taken as new variables to give the desired canonical form

$$(2) \qquad \eta_i' = \lambda_i\,\eta_i \qquad (i = 1, \cdots, r).$$

Denote its matrix $B^{-1}\,N_2\,B$ by P_2. Similarly, denote $B^{-1}\,N_j\,B$ by P_j. Hence our matrix N (and thus M) has been reduced to

$$P = x_1\,I + x_2\,P_2 + \cdots + x_n\,P_n.$$

For $n > 2$, the further normalization of P is to be accomplished by means of a matrix K such that $K^{-1}\,P_2\,K = P_2$. But the only linear transformation commutative with (2), in which $\lambda_1, \cdots, \lambda_r$ are distinct, is seen at once to be

$$(3) \qquad \eta_i' = k_i\,\eta_i \qquad (i = 1, \cdots, r),$$

where the k's need not be distinct, but each is $\neq 0$. If

$$\eta_i' = \sum_{j=1}^{r} c_{ij}\,\eta_j \qquad (i = 1, \cdots, r)$$

is the transformation whose matrix is P_3, and if K is the matrix of (3), then, in view of the above interpretation, $K^{-1}\,P_3\,K$ is the matrix of

$$\zeta_i' = \sum_{j=1}^{r} \frac{k_i}{k_j} c_{ij}\,\zeta_j \qquad (i = 1, \cdots, r;\ \zeta_i = k_i\,\eta_i).$$

Thus the maximum simplification possible in $P_3, \cdots, P_n$ is to make $r - 1$ non-vanishing elements take the value unity, so that at most $(n-2)r^2 - (r-1)$ parameters appear in their canonical forms. Taking account also of $\lambda_1, \cdots, \lambda_r$ and of the factor initially removed from D, we conclude

* *Ibid.*, p. 222.

that D can be given a form containing at most $(n-2)r^2+2$ parameters. A generalization of Theorem 1 is given in § 10.

3. The number of terms in the general form of degree r in n variables is known to be the binomial coefficient

$$\binom{r+n-1}{r} = \frac{(r+n-1)!}{r!(n-1)!}.$$

This follows by two-fold induction since in

$$\binom{r+n-2}{r-1} + \binom{r+n-2}{r} = \binom{r+n-1}{r},$$

the first symbol therefore enumerates the terms with the factor x_n and the second symbol enumerates the terms lacking x_n.

4. By §§ 2–3, the general form of degree r in n variables $(n > 2)$ is not expressible in determinantal form if

$$\binom{r+n-1}{r} > (n-2)r^2+2. \tag{4}$$

If $r = 2$, this condition reduces to $(n-3)(n-4) > 0$. If $r = 3$, it is

$$n^3 + 3n^2 - 52n + 96 > 0$$

and holds when $n \geqq 5$. If $r \geqq 4$ and $n \geqq 7$, we have

$$\binom{r+n-1}{n-1} > \frac{(n+3)(n+2)(n+1)n\cdots 7(r+2)(r+1)}{(n-1)!},$$

since we have replaced r by 4 in all but the last two factors. The factors $n-1, n-2, \cdots, 7$ (which are absent if $n = 7$) may be cancelled. We get

$$(n+3)(n+2)(n+1)n(r+2)(r+1)/6!,$$

which will exceed $t = (n-2)r^2+2$, since $(n-2)(r+2)(r+1) > t$, provided

$$(n+3)(n+2)(n+1)n/6! \geqq n-2.$$

The latter may be written in the form

$$(n-7)(n^3+13n^2+102n)+1440 \geqq 0.$$

It remains to treat the cases $n \leqq 6$. For $n = 3$, (4) fails if $r \geqq 2$, since $\binom{r+2}{2} \leqq r^2+2$ for $(r-1)(r-2) \geqq 0$. But for $n = 4$, (4) holds if $(r-1)(r-2)(r-3) > 0$. For $n = 5$ and $n = 6$, (4) becomes respectively

$$r^4 + 10r^3 - 37r^2 + 50r - 24 > 0,$$

$$r^5 + 15r^4 + 85r^3 - 255r^2 + 274r - 120 > 0,$$

each of which evidently holds if $r \geqq 3$. Hence we have

THEOREM 2. *The general form of degree r in n variables, $n > 2$, is not expressible in determinantal form if $r = 2$ or 3, $n > 4$, and if $r \geqq 4$, $n \geqq 4$, and hence unless $n = 3$, r any, or $n = 4$, $r = 2$ or 3.*

5. Since general quadric and cubic surfaces are expressible in determinantal form (§ 1), there remains only the case $n = 3$.

For $n = r = 3$, we may employ the canonical forms,* omitting those with a linear factor (§ 6):

$$x^3 + y^3 + z^3 - mxyz \equiv \begin{vmatrix} x & y & z \\ z & x & ay \\ a^{-1}y & z & x \end{vmatrix}, \qquad a + \frac{1}{a} = m - 1,$$

$$x^3 + y^3 - xyz \equiv \begin{vmatrix} x & y & z \\ 0 & x & y \\ y & 0 & x \end{vmatrix}, \qquad x^3 + yz^2 \equiv \begin{vmatrix} x & y & 0 \\ 0 & x & z \\ z & 0 & x \end{vmatrix}.$$

I obtained the determinants by inspection.

For $n = 3$, $r = 4$, a general quartic curve can be given the form

$$a^2 b^2 + c^2 d^2 + e^2 f^2 - 2abcd - 2abef - 2cdef = 0,$$

where a and b, c and d, e and f are three pairs of bitangents of a Steiner set.† I find that this function equals the determinant

$$\begin{vmatrix} a & 0 & c & f \\ 0 & a & e & d \\ d & f & b & 0 \\ e & c & 0 & b \end{vmatrix}.$$

But the actual determination of the bitangents depends upon the solution of an equation of very high order. Also numerous radicals appear in the expressions for m and the canonical variables in the above cubic form. The method next explained is not only general, but employs no irrationals other than the roots $\lambda_1, \cdots, \lambda_r$.

6. The following method enables us to express the equation $f = 0$ of any plane curve of order r as a determinant of order r whose elements are linear functions of x, y, z. It will suffice to prove this for irreducible forms f. For, if $f = f_1 f_2$, where f_i is of degree r_i and is expressible as a determinant of order r_i of matrix M_i, then f equals the determinant of the matrix

$$\begin{pmatrix} M_1 & O \\ O & M_2 \end{pmatrix},$$

where O is a matrix all of whose elements are zero.

* P. Gordan, these Transactions, vol. 1 (1900), p. 402.

† Miller, Blichfeldt and Dickson, *Finite Groups*, 1916, p. 355.

We suppose merely that f has no repeated factor. Then there exists a straight line which cuts the curve in r distinct points.[*] Take it as the side $z = 0$ of a triangle of reference. Take as the side $y = 0$ any line not meeting $z = 0$ at one of its r intersections with the curve. Then $(1, 0, 0)$ is not on the curve, and the coefficient of x^3 in f may be assumed to be unity. Hence, for $z = 0$, f reduces to a product $X_1 X_2 \cdots X_r$ of r distinct linear functions $X_i = x + \lambda_i y$. Thus

$$f = X_1 X_2 \cdots X_r + \sum_{k=1}^{r} z^k F_k(y, x), \tag{5}$$

where F_k is a binary form of order $r - k$.

We shall prove that every such form f, in which $\lambda_1, \cdots, \lambda_r$ are distinct, can be expressed as a determinant of the type suggested by § 2:

$$\begin{vmatrix} X_1 + c_{11} z & c_{12} z & \cdots & c_{1r} z \\ \cdot & \cdot & \cdot & \cdot \\ c_{r1} z & c_{r2} z & \cdots & X_r + c_{rr} z \end{vmatrix}. \tag{6}$$

There are $\frac{1}{2}(r + 2)(r + 1)$ coefficients in a general ternary form of order r. In (5) we have identified $f(x, y, 0)$ with $X_1 \cdots X_r$, thus fixing $r + 1$ coefficients; there remain $\frac{1}{2}(r^2 + r)$ coefficients. Hence the identification of (6) with (5) involves as many conditions as there are c_{ij} in and below the main diagonal. Accordingly we shall assign simple values to the remaining c_{ij}:

$$c_{i\,i+1} = 1, \qquad c_{ij} = 0 \quad (j > i + 1;\ i, j = 1, \cdots, r). \tag{7}$$

This choice is in accord with the general theory in § 2, where it was shown that, without altering determinant (6), we may assign the value unity to $r - 1$ non-vanishing c's.

We proceed to prove that the c_{ij} $(j \leqq i)$ can be uniquely determined so that determinant (6), subject to (7), becomes identical with any given form (5). Use is made of the known expansion of an "axial" determinant (6). First, the terms linear in z are

$$\sum_{i=1}^{r} c_{ii}\, z X_1 \cdots X_{i-1} X_{i+1} \cdots X_r .$$

This sum will be identical with $zF_1(y, x)$, where F_1 is any given binary form

[*] For, if every line $z = rx + sy$ cuts $f = 0$ in points two of which coincide,

$$F(x, y) \equiv f(x, y, rx + sy) = 0$$

has a double root for every r, s, whence

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + r\frac{\partial f}{\partial z} = 0, \qquad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} + s\frac{\partial f}{\partial z} = 0$$

for every r, s. Thus every point on $f = 0$ is a singular (multiple) point.

of order $r-1$, if they are equal for the r values $x=-\lambda_i y$ $(i=1,\cdots,r)$ for which the X_i vanish. The resulting conditions

$$c_{ii}(\lambda_1-\lambda_i)\cdots(\lambda_{i-1}-\lambda_i)(\lambda_{i+1}-\lambda_i)\cdots(\lambda_r-\lambda_i)=F_1(1,-\lambda_i)$$

uniquely determine c_{ii} $(i=1,\cdots,r)$. Next, the terms of (6) quadratic in z are

$$\sum\begin{vmatrix} c_{ii} & c_{ij} \\ c_{ji} & c_{jj}\end{vmatrix} z^2 \frac{X_1\cdots X_r}{X_i X_j} \qquad (i,j=1,\cdots,r;\ i<j).$$

This sum is to be identified with $z^2F_2(y,x)$. If $j>i+1$, then $c_{ij}=0$ and the two-rowed determinant equals the previously determined number $c_{ii}c_{jj}$. If $j=i+1$, then $c_{ij}=1$ and the diagonal term is known. Transposing all the known terms and combining them with z^2F_2, we are to identify

$$-\sum_{i=1}^{r-1} c_{i+1,\,i}X_1\cdots X_{i-1}X_{i+2}\cdots X_r$$

with a known binary form of order $r-2$. It suffices* for this purpose to take in turn $X_1=0$, $X_2=0$, $\cdots$, $X_{r-1}=0$. Of the resulting conditions, the first determines c_{21}, the second determines c_{32} in terms of c_{21}, ..., the last determines $c_{r,\,r-1}$ in terms of $c_{r-1,\,r-2}$, so that they uniquely determine all of the $c_{i+1,\,i}$.

To make the general step of the proof by induction, we assume that the terms of (6) of degrees $1,2,\cdots,k-1$ in z have been identified with the terms of (5) of corresponding degrees by the unique determination of the c_{ii}, $c_{i+1,\,i}$, $\cdots$, $c_{i+k-2,\,i}$. The terms of (6) of degree k in z are

$$\sum\begin{vmatrix} c_{i_1i_1} & c_{i_1i_2} & \cdots & c_{i_1i_k} \\ \cdot & \cdot & \cdot & \cdot \\ c_{i_ki_1} & c_{i_ki_2} & \cdots & c_{i_ki_k}\end{vmatrix} z^k\frac{X_1\cdots X_r}{X_{i_1}\cdots X_{i_k}} \quad \begin{pmatrix} i_1,\cdots,i_k=1,\cdots,r; \\ i_1<i_2<\cdots<i_k\end{pmatrix}.$$

This sum is to be identified with $z^kF_k(y,x)$. If in the determinant just written every element just above the main diagonal is unity, then

$$i_2=i_1+1,\quad i_3=i_2+1=i_1+2,\quad \cdots,\quad i_k=i_{k-1}+1=i_1+k-1,$$

and the difference between the subscripts of any c other than $c_{i_ki_1}$ is $\leqq k-2$, so that the c is one of those previously determined, while the minor of the exceptional c equals unity. We shall prove that the expansions of the remaining determinants D involve only previously determined c's. Then after transposing known terms and combining them with z^kF_k, we have left only the product of z^k by

$$\sum_{i=1}^{r-k+1}(-1)^{k-1}c_{i+k-1,\,i}\frac{X_1\cdots X_r}{X_iX_{i+1}\cdots X_{i+k-1}},$$

* Or we may determine $c_{r,\,r-1}$ from $X_r=0$ and work back half way or all the way.

which is to be identified with a given binary form of order $r - k$. The conditions obtained by taking in turn $X_1 = 0, \cdots, X_{r-k+1} = 0$ determine these c's in turn. Hence the induction is complete.

It remains to prove the statement regarding any determinant D which has at least one zero element c just above the main diagonal. Then, by (7), zero is the value of every element of D which lies in the rectangle bounded by c and the elements above it and the elements to the right of it. Hence D is the product of two principal minors (each having its diagonal elements on the main diagonal of D). In case either minor has a zero element just above its diagonal, it decomposes similarly into a product of principal minors. Hence D is a product of two or more principal minors each of which has either a single element c_{ii} or is a principal minor P all of whose t elements just above the diagonal are equal to unity. If the diagonal elements of such a $(t+1)$-rowed P are

$$c_{ii} \quad (i = i_l, i_m, i_n, \cdots, i_w),$$

then, by (7),

$$i_m = i_l + 1, \qquad i_n = i_m + 1 = i_l + 2, \qquad \cdots, \qquad i_w = i_l + t,$$

so that the maximum difference of subscripts in any element of P is $i_w - i_l = t$. Since P has fewer rows than D, $t + 1 < k$, and $t \leqq k - 2$. Hence by the hypothesis for the induction, every element of P is among those previously determined.

THEOREM 3. *Every plane curve can be represented by equating to zero a determinant whose elements are linear functions. In particular, any ternary form without a repeated factor can be transformed linearly into mf, where m is a constant and $f(x, y, 0) = X_1 \cdots X_r$, $X_i = x + \lambda_i y$, $\lambda_1, \cdots, \lambda_r$ being distinct. Then f can be expressed in one and but one way as a determinant*

$$\begin{vmatrix} X_1 + c_{11} z & z & 0 & \cdots & 0 \\ c_{21} z & X_2 + c_{22} z & z & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{r1} z & c_{r2} z & c_{r3} z & \cdots & X_r + c_{rr} z \end{vmatrix},$$

in which the elements just above the main diagonal equal z and the remaining elements above the diagonal equal zero.

7. The problem for quadratic forms may be treated with attention to rationality. A determinant of order 2 or 3, whose elements are linear forms in x, y, z, w with coefficients in a given field F (or domain of rationality) evidently vanishes for a set of values, not all zero, in F. Hence a quadric surface which is representable as such a determinant must have a point with coördinates in F, which may be taken to be $(1, 0, 0, 0)$. If the surface is not a cone, the coefficient of x may be taken as the new variable y. Then by

adding to x a suitable linear function of y, z, w, we obtain $xy + Q(z, w)$, where Q is a binary quadratic form with coefficients in F. A determinant representing it can evidently be given the form

$$\begin{vmatrix} x + A & B \\ C & y \end{vmatrix},$$

where A is free of x, while B and C are free of x and y. Thus $A \equiv 0$, and the condition is that Q have linear factors with coefficients in F, so that the tangent plane $y = 0$ at $(1, 0, 0, 0)$ cuts the surface in rational lines. For a cone or conic, we note that a ternary quadratic form which lacks x^2 can be transformed rationally* into $xy + az^2$ or $Q(y, z)$. But every binary form can be expressed rationally in determinantal form:

$$a_0 x^r + a_1 x^{r-1} y + \cdots + a_r y^r = \begin{vmatrix} a_0 x + a_1 y & y & 0 & 0 & \cdots & 0 \\ -a_2 y & x & y & 0 & \cdots & 0 \\ a_3 y & 0 & x & y & \cdots & 0 \\ -a_4 y & 0 & 0 & x & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (-1)^r a_{r-1} y & 0 & 0 & 0 & \cdots & y \\ (-1)^{r+1} a_r y & 0 & 0 & 0 & \cdots & x \end{vmatrix}.$$

THEOREM 4. *Every binary form can be expressed rationally in determinantal form. A rational quadric surface not a cone can be expressed rationally in determinantal form if and only if it has a rational point the tangent plane at which cuts the surface in rational lines. A conic (or quadric cone) can be represented rationally in determinantal form if and only if it has a rational point (not the vertex).*

If we desire to ignore irrationalities, we take F to be the field of all complex numbers. A summary of our results gives

THEOREM 5. *Every binary form, every ternary form, every quaternary quadratic form, and a sufficiently general quaternary cubic form can be expressed in determinantal form. No further general form has this property.*

I have treated elsewhere† the problem of quaternary cubic forms with attention to rationality. The number 20 of coefficients equals the number of disposable parameters in the determinant (§ 2), and the problem depends in general upon the solution of a single algebraic equation. The notations suggested by § 2 are less convenient for this problem than those used in the paper cited.

8. We shall examine briefly the conditions under which a given rational ternary cubic form T is expressible rationally in determinant form. If T

* Dickson, *Algebraic Invariants*, 1914, p. 24.

† American Journal of Mathematics, April, 1921.

vanishes at a rational point, the problem has been treated fully in the paper last cited. In the contrary case, we may assume that the coefficient of x^3 is unity and that the terms in $x^2 y$ and $x^2 z$ are lacking. The matrix of the determinant may be taken to be $x_1 I + x_2 N_2 + zN_3$, where $x_1 = x$, $x_2 = y$. Since T is not zero at a rational point, $|x_1 I + x_2 N_2| \neq 0$ when x_1 and x_2 are rational and not both zero. As shown in § 2, the characteristic equation of N_2 therefore has no rational root, and hence has a single invariant factor $\lambda^3 - \alpha\lambda - \beta$, so that there exists (§ 9) a matrix of rational coefficients which transforms N_2 into P_2 and N_3 into* P_3:

$$P_2 = \begin{pmatrix} 0 & \alpha & \beta \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad P_3 = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}.$$

The determinant of $xI + yP_2 + zP_3$ will be identical with

$$T = x^3 - \alpha xy^2 + \beta y^3 + Bxyz + Cy^2 z + Dxz^2 + Eyz^2 + Fz^3$$

if and only if

$$a + e + k = 0, \qquad b + \alpha d + \beta g + f = -B, \qquad c - \alpha k + \beta d + \beta h = C,$$

$$\begin{vmatrix} a & b \\ d & e \end{vmatrix} + \begin{vmatrix} a & c \\ g & k \end{vmatrix} + \begin{vmatrix} e & f \\ h & k \end{vmatrix} = D,$$

$$-\begin{vmatrix} b & c \\ h & k \end{vmatrix} - \alpha \begin{vmatrix} d & f \\ g & k \end{vmatrix} - \begin{vmatrix} a & c \\ d & f \end{vmatrix} + \beta \begin{vmatrix} d & e \\ g & h \end{vmatrix} = E, \qquad |P_3| = F.$$

Employ the linear equations to express a, b, c in terms of the remaining letters. In D, g and f enter linearly, the latter with the coefficient $d - h$. If $h = d$, the result of substituting the values of g and f given by our equations (D) and (E) into $|P_3| = F$ is of the ninth degree in each d and e and of the eighth degree in k. Since this special case thus involves hopeless difficulties, we take $h = d + m$, $m \neq 0$. Then equations (D), (E) and (F) give

$$f = \frac{g\gamma - \delta}{m}, \qquad \gamma \equiv m\beta - \alpha k + 3\beta d - C,$$

$$\delta \equiv D + e^2 + ek + k^2 - Bd - \alpha d^2,$$

$$\alpha\gamma g^2 + \{\gamma(e + 2k) + m\beta(k - e) - \alpha\delta\}g - m\epsilon - \delta(e + 2k) = 0,$$

$$\epsilon \equiv E - Bk - C(m + 2d) - 2\alpha dk - \alpha mk + 3\beta d^2 + 3\beta md + \beta m^2,$$

* For the normalization of P_3 we have available the most general matrix

$$rI + sP_2 + t\begin{pmatrix} 0 & \beta & 0 \\ 0 & 0 & \beta \\ 1 & 0 & -\alpha \end{pmatrix}$$

commutative with P_2. But it transforms P_3 into such a complicated matrix that normalization does not seem worth while.

$$m^2 F = \begin{vmatrix} -e-k & m(-B-\alpha d-\beta g)-g\gamma+\delta & C-m\beta-2\beta d+\alpha k \\ md & m^2 e & g\gamma-\delta \\ g & m^2+md & k \end{vmatrix}.$$

The resultant of our quadratic and cubic equations in g is readily seen to be of the sixth degree in m, with coefficients involving d, e, k to higher degree. Even in the simple case in which T has a rational point, the problem depends upon a cubic irrationality.

9. Two pairs of bilinear forms with matrices M, N and M', N' with elements in a field F and such that N and N' are non-singular (determinant $\neq 0$) are called equivalent with respect to F if there exist non-singular matrices P and Q with elements in F such that $PMQ = M'$, $PNQ = N'$. Necessary and sufficient conditions for equivalence are known to be the identity of the invariant factors of $M - \lambda N$ with those of $M' - \lambda N'$. The special case $N = N' = I$ shows that matrix (or substitution) M can be transformed within F into $Q^{-1} MQ = M'$ if and only if the invariant factors of the characteristic determinants $|M - \lambda I|$ and $|M' - \lambda I|$ of M and M' are the same. This case finds application in many branches of mathematics.

The last result has been proved independently* of the theory of pairs of bilinear forms. From it we can deduce the latter theory. First, let M, N be equivalent to M', N', so that $MN^{-1} = J$, I are equivalent to $M' N'^{-1} = J'$, I, whence $AJB = J'$, $AIB = I$, $A = B^{-1}$. Thus B transforms J into J' and their characteristic determinants $|MN^{-1} - \lambda I|$ and $|M' N'^{-1} - \lambda I|$ have the same invariant factors. The latter are not altered when the determinants are multiplied by the constants $|N|$ and $|N'|$, respectively. Hence $|M - \lambda N|$ and $|M' - \lambda N'|$ have the same invariant factors. This necessary condition is also a sufficient condition for equivalence of the pairs. For, we then have $B^{-1} JB = J'$, $J \equiv MN^{-1}$, $J' \equiv M' N'^{-1}$, where the elements of B are in F. Then, if λ is arbitrary,

$$B^{-1}(MN^{-1} - \lambda I)B = M' N'^{-1} - \lambda I,$$

$$B^{-1}(M - \lambda N)N^{-1}B = (M' - \lambda N')N'^{-1},$$

$$B^{-1}(M - \lambda N)C = M' - \lambda N', \qquad C = N^{-1}BN'.$$

Thus B^{-1} serves as pre-factor and C as post-factor to convert M into M' and N into N', so that the pairs are equivalent.

The ideas in this paper evidently apply also to the question of the equivalence of n-tuples of bilinear forms, and are being developed by one of my students.

10. A generalization of Theorem 1 is given by

THEOREM 6. *Any r-rowed determinant D whose elements are linear homo-*

* Dickson, these Transactions, vol. 3 (1902), pp. 290–2.

geneous functions of $x_1, \cdots, x_n$ can be expressed in a canonical form involving not more than $(n-2)r^2+2$ parameters when $n>2$.

The matrix of D is of the form $M = x_1 M_1 + \cdots + x_n M_n$.

Since we may assume that D is not identically zero, there exist constants $c_1, \cdots, c_n$ such that the determinant of $C = c_1 M_1 + \cdots + c_n M_n$ is not zero. Replacing $x_1, \cdots, x_n$ by

$$c_1 x_1 + d_1 x_2 + \cdots + k_1 x_n, \quad \cdots, \quad c_n x_1 + d_n x_2 + \cdots + k_n x_n,$$

where the $d_i, \cdots, k_i$ are chosen so that the determinant of these linear forms is not zero, we see that M is replaced by $x_1 C + \cdots$. Returning to the initial notations, we may therefore assume that $|M_1| \neq 0$.

Proceeding as in § 2, we reduce M to an equivalent matrix

$$P = x_1 I + x_2 P_2 + \cdots + x_n P_n,$$

in which P_2 is the canonical form of a matrix (or linear substitution) obtained by transformation. We proved our theorem in § 2 for the case in which P_2 is of the special form (2) with $\lambda_1, \cdots, \lambda_r$ distinct. In case equalities occur among these λ's, transformations (3) are included among those commutative with P_2, so that we can accomplish the same (and further) specialization of the parameters in $P_3, \cdots, P_n$ as in § 2. Hence for two reasons the number of parameters in $P_2, \cdots, P_n$ is smaller than before.

Finally, let the canonical P_2 be of its most general form:*

$$(8)\quad \begin{gathered} \eta_1' = \lambda_1 \eta_1 + \eta_2, \quad \eta_2' = \lambda_1 \eta_2 + \eta_3, \quad \cdots, \quad \eta_{e_1-1}' = \lambda_1 \eta_{e_1-1} + \eta_{e_1}, \\ \eta_{e_1}' = \lambda_1 \eta_{e_1}; \end{gathered}$$

$$(9)\quad \begin{gathered} \zeta_1' = \lambda_2 \zeta_1 + \zeta_2, \quad \zeta_2' = \lambda_2 \zeta_2 + \zeta_3, \quad \cdots, \quad \zeta_{e_2-1}' = \lambda_2 \zeta_{e_2-1} + \zeta_{e_2}, \\ \zeta_{e_2}' = \lambda_2 \zeta_{e_2}; \end{gathered}$$

etc. Now (8) is commutative with†

$$(10)\quad \eta_i' = k_1 \eta_i + k_2 \eta_{i+1} + \cdots + k_{e_1-i+1} \eta_{e_1} \qquad (i = 1, \cdots, e_1).$$

To verify this fact, we may take $\lambda_1 = 0$ in (8), since we may subtract $\lambda_1 I$ from the matrix of (8); then the product of the modified (8) and (10), in either order, replaces η_i by

$$k_1 \eta_{i+1} + k_2 \eta_{i+2} + \cdots + k_{e_1-i} \eta_{e_1}.$$

Similarly, (9) is commutative with

$$(11)\quad \zeta_i' = l_1 \zeta_i + l_2 \zeta_{i+1} + \cdots + l_{e_2-i+1} \zeta_{e_2} \qquad (i = 1, \cdots, e_2).$$

* Dickson, *Linear Groups*, p. 223; Bôcher, *Higher Algebra*, p. 293.

† We do not need the fact that every linear transformation on $\eta_1, \cdots, \eta_{e_1}$ which is commutative with (8) is of the form (10). If the λ's are all distinct, every transformation commutative with P_2 is given by (10), (11), etc.

Transformations (10), (11), etc. are available for the normalization of $P_j\ (j > 2)$. Let P_j replace η_{e_1} by

$$t_1\,\eta_1 + \cdots + t_{e_1}\,\eta_{e_1} + f, \tag{12}$$

where f is a linear function of the ζ's, etc. If K is the matrix of the transformation defined by (10), (11), etc., $K^{-1}\,P_j\,K$ replaces η_{e_1} by the product of k_1^{-1} by the function by which (10), (11), etc., replace (12), the product being

$$\frac{1}{k_1}\left\{f' + \sum_{i=1}^{e_1} t_i\,(k_1\,\eta_i + k_2\,\eta_{i+1} + \cdots + k_{e_1-i+1}\,\eta_{e_1})\right\}$$

$$= t_1\,\eta_1 + \left(t_2 + \frac{k_2}{k_1}t_1\right)\eta_2 + \left(t_3 + \frac{k_2}{k_1}t_2 + \frac{k_3}{k_1}t_1\right)\eta_3 + \cdots$$

$$+ \left(t_{e_1} + \frac{k_2}{k_1}t_{e_1-1} + \cdots + \frac{k_{e_1}}{k_1}t_1\right)\eta_{e_1} + \frac{f'}{k_1},$$

where f' is the function by which (11), etc. replace f. Hence the ratios

$$k_2/k_1,\ \cdots,\ k_{e_1}/k_1,\ l_1/k_1,\ l_2/l_1,\ \cdots,\ l_{e_2}/l_1,\ \cdots$$

are available to specialize as many parameters in P_j. Their number is $e_1 + e_2 + \cdots + e_t - 1$ if $e_1 + \cdots + e_t = r$. In § 2 we could specialize only this number $r - 1$ of parameters in the $P_j\ (j > 2)$. We now have only t distinct λ's instead of the former r. Hence the total number of parameters found in § 2 is the true maximum.

THE UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

PSEUDO-CANONICAL FORMS AND INVARIANTS OF SYSTEMS OF PARTIAL DIFFERENTIAL EQUATIONS*

BY

ALFRED L. NELSON

1. Introduction

The theory of invariants of the linear ordinary differential equation dates from 1862, when Sir James Cockle† obtained certain seminvariants. In later papers he made progress in the theory, while other writers,‡ notably Laguerre, Brioschi, Halphen and Forsyth, extended his results. The theory thus developed presents a striking analogy to the theory of algebraic invariants. One point of similarity, with which this paper has much to do, is the fact that by means of a transformation of the type $y = \lambda(x)\bar{y}$, the linear differential equation

$$y^{(n)} + p_1 ny^{(n-1)} + \cdots + p_n y = 0$$

may be made to assume a canonical form, in which the coefficient of $y^{(n-1)}$ is zero. The remaining coefficients are then unchanged when the equation is subjected to any transformation of the type in question, i.e., they are absolute seminvariants. The seminvariants thus obtained are moreover a fundamental set, in the sense that any seminvariant whatever is a function of them and of their derivatives.

The work of Wilczynski§ has shown how the previous theory may be extended to include completely integrable systems of linear homogeneous partial

* Presented to the Society, September 2, 1919.

† *The Correlations of Analysis*, Philosophical Magazine, ser. 4, vol. 24, pp. 531–534.

‡ In the first pages of an article "*Invariants of the General Linear Differential Equation and their Relation to the Theory of Continuous Groups*," American Journal of Mathematics, vol. 21 (1899), pp. 25–84, Bouton gives a brief sketch of these contributions.

§ *Projective Differential Geometry of Curved Surfaces*, these Transactions, vol. 8 (1907), pp. 233–260.

One-Parameter Families and Nets of Plane Curves, these Transactions, vol. 12 (1911), pp. 473–510.

Sur la Théorie Générale des Congruences, Mémoires de l'Académie Royale de Belgique, Classe des Sciences, ser. 2, vol. 3 (1910–1912).

differential equations. In various papers, he and others* have obtained canonical forms of certain such systems, making use of the transformation $y = \lambda(u, v)\bar{y}$. As in the theories referred to, the coefficients of a canonical form are a fundamental set of seminvariants. Green† has proved that for a very large class of completely integrable systems, such a canonical form can always be obtained, and that its coefficients will have the usual properties. Finally should be mentioned a paper of Wilczynski,‡ in which is given a general proof covering all known theories of invariants, as well as many others.

It is characteristic of all the work just referred to that the actual existence of a canonical form has been considered necessary to the proof that the coefficients of such a form are seminvariants. This has resulted, in so far as partial differential equations are concerned, in a failure to obtain canonical forms characterized by the vanishing of certain *coefficients*.§ Instead, certain *functions of the coefficients* have been made to vanish. The seminvariants secured, therefore, have not been as simple as might be desired. The author || has shown that, for a large class of completely integrable systems, a fundamental set of simpler seminvariants can be constructed as the coefficients of a *pseudo-canonical form*. These non-existent pseudo-canonical forms are characterized by the vanishing of certain coefficients. The proof used seemed more complicated than necessary, and one of the aims of the present paper is to furnish a simpler one.

A second purpose is the extension of the theory presented in the previous article to include semi-covariants, invariants and covariants.

Attention will be confined to partial differential equations with a single dependent variable, although the facts exhibited will hold for systems having two or more dependent variables, provided the transformations used are of the type discussed in this paper. It must be admitted, however, that such

* W. W. Denton, *Projective Differential Geometry of Developable Surfaces*, these Transactions, vol. 14 (1913), pp. 175–208.

G. M. Green, *Projective Differential Geometry of Triple Systems of Surfaces*, Columbia Dissertation, 1913.

G. M. Green, *Projective Differential Geometry of One-Parameter Families of Space Curves, and Conjugate Systems of Curves on a Surface*, American Journal of Mathematics, vol. 37 (1915), pp. 215–246.

† *Linear Dependence of Functions of Several Variables, and Completely Integrable Systems of Homogeneous Linear Partial Differential Equations*, these Transactions, vol. 17 (1916), pp. 483–516.

‡ *Invariants and Canonical Forms*, Proceedings of the National Academy of Sciences, vol. 4 (1918), pp. 300–305.

§ The one exception to this statement is the theory of curved surfaces, discussed by Wilczynski in the paper referred to.

|| *Note on Seminvariants of Systems of Partial Differential Equations*, American Journal of Mathematics, vol. 41 (1919), pp. 123–132.

transformations (in cases which involve two dependent variables) are not always the ones of greatest interest geometrically.*

Ordinary differential equations will not be considered, for the reason that the methods of this paper, insofar as they apply to ordinary differential equations, reduce to old methods in such cases.

Particular attention must be called to a change in terminology from the author's paper cited. What were there called *pseudo-canonical forms* are now termed *pseudo-semi-canonical forms,* the first name being reserved for certain new unique forms.

The author wishes to thank Professor Wilczynski for valuable criticisms, which have resulted in an improvement in the form of this paper. He also desires to make the following further acknowledgments: Dr. A. L. Miller, in setting up a completely integrable system of equations to be used as a basis for a special theory, accidentally discovered that seminvariant coefficients resulted in that special case when a certain unallowable transformation was employed. Acting upon this hint, the author succeeded in establishing the results contained in his previous paper. While considering the feasibility of the extensions embraced in the present paper, special results of Dr. W. W. Denton in the theory of developable surfaces were of assistance.

2. Pseudo-semi-canonical forms and seminvariants

Let us assume that the completely integrable system of differential equations (a) has one dependent variable, y, and n independent variables, $u_1, \cdots, u_n$; that it consists of p equations, each of which expresses a certain derivative of y linearly and homogeneously in terms of certain q primary derivatives (including y itself); that no primary derivative is of higher order than any of the left members of the equations (a); that if a given y-derivative,

$$\frac{\partial^{l_1+\cdots+l_n} y}{\partial u_1^{l_1} \cdots \partial u_n^{l_n}},$$

occurs in any equation of (a), then all derivatives of lower order, from which the given derivative may be obtained by differentiation, are also present in that equation.†

We shall group the various y-derivatives and coefficients of (a) as follows:

* See, for example, Wilczynski's basis for the study of congruences of straight lines, in his prize memoir, "*Sur la Théorie Générale des Congruences,*" l. c. Compare with that used by Green, in section 8 of his paper, "*Projective Differential Geometry of One-parameter Families of Space Curves etc.,*" l. c. The former employs the transformations used in the present paper, while the latter does not.

† The purpose of this assumption is to insure that a transformation of the dependent variable shall replace (a) by a system of exactly the same form. Cf. Green, *Linear Dependence of Functions of Several Variables*, etc., l. c., section 6.

A y-derivative of the same order as the left member of its equation is said to be of the zeroth class. A y-derivative of order i less than the left member of its equation is of the ith class. A coefficient of an ith class y-derivative is of the ith class.

The equation

$$(1a) \qquad y = \lambda(u_1, \cdots, u_n)\bar{y}$$

yields by differentiation

$$y_{u_i} = \lambda \bar{y}_{u_i} + \lambda_{u_i}\bar{y} \qquad (i = 1, \cdots, n),$$

$$(1b) \qquad \cdots\cdots\cdots\cdots\cdots\cdots$$

$$y_{l_1,\cdots,l_n} = \sum \binom{l_1}{p_1}\cdots\binom{l_n}{p_n}\lambda_{p_1,\cdots,p_n}\bar{y}_{l_1-p_1,\cdots,l_n-p_n},$$

$p_1 = 0, \cdots, l_1, p_2 = 0, \cdots, l_2, \cdots, p_n = 0, \cdots, l_n$, where

$$y_{l_1,\cdots,l_n} = \frac{\partial^{l_1+\cdots+l_n} y}{\partial u_1^{l_1}\cdots \partial u_n^{l_n}}, \qquad \binom{l_i}{p_i} = \frac{l_i!}{(l_i - p_i)!\,p_i!}.$$

When equations (1) are substituted in the system (a) there results a new system, $(\bar{a})$, of the same form as (a), in which the independent variables are, as before, $u_1 . \cdots . u_n$, while the dependent variables are $\bar{y}, \bar{y}_{u_i}$ $(i = 1, \cdots, n)$, etc. These new variables may be expressed explicitly in terms of y, y_{u_i}, etc., as follows:

$$(2a) \qquad \bar{y} = \mu y, \qquad \mu = 1/\lambda,$$

$$\bar{y}_{u_i} = \mu\left[y_{u_i} + \frac{\mu_{u_i}}{\mu}\cdot y\right] \qquad (i = 1, \cdots, n),$$

$$(2b) \qquad \cdots\cdots\cdots\cdots\cdots\cdots,$$

$$\bar{y}_{l_1,\cdots,l_n} = \mu\sum\binom{l_1}{p_1}\cdots\binom{l_n}{p_n}\frac{\mu_{p_1,\cdots,p_n}}{\mu}\cdot y_{l_1-p_1,\cdots,l_n-p_n},$$

with $p_1, p_2, \cdots, p_n$ as in (1b).

A coefficient of $(\bar{a})$ of the first class is of the type

$$(3) \qquad \bar{\beta} = \beta + \sum_{i=1}^{n} p_i\,\alpha_i\frac{\mu_{u_i}}{\mu},$$

where the p_i are integers (including zero) and the α_i are seminvariant zeroth class coefficients of (a).*

Let us take† n equations of the type $\bar{\beta} = 0$, choosing the coefficients β in such a way as to permit the solution of this set of n equations for μ_{u_i}/μ $(i = 1,$

* Cf. A. L. Nelson, l. c., section 2. It was there tacitly assumed that a first-class coefficient would be changed (if at all), by the transformation (1a), by the addition of a multiple of a single such μ-fraction. The structure of higher class coefficients of $(\bar{a})$ is also discussed in this section.

† The possibility of this will be discussed in section 8.

$\cdots , n)$, regarding these fractions as independent unknowns. We denote these solutions by (μ_{u_i}/μ).

It is of course impossible to find a function, $\mu(u_1, \cdots , u_n)$, which will satisfy the n differential equations

$$\bar{\beta}_i = 0, \qquad (i = 1, \cdots , n),$$

unless certain integrability conditions are satisfied. These conditions are

$$\frac{\partial}{\partial u_i}\left(\frac{\mu_{u_j}}{\mu}\right) = \frac{\partial}{\partial u_j}\left(\frac{\mu_{u_i}}{\mu}\right) \qquad (i, j = 1, \cdots , n). \tag{4}$$

Nevertheless, without assuming that equations (4) hold, we substitute the solutions (μ_{u_i}/μ) for μ_{u_i}/μ in all the coefficients of $(\bar{a})$.

However, since (μ_{u_i}/μ) and (μ_{u_j}/μ) $(i, j = 1, \cdots , n;\ i \neq j)$, are two distinct functions of the coefficients of (a), there are two independent ways of substituting for $\mu_{u_i u_j}/\mu$, namely,

$$\frac{\mu_{u_i u_j}}{\mu} = \left(\frac{\mu_{u_i}}{\mu}\right)_{u_j} + \left(\frac{\mu_{u_i}}{\mu}\right)\left(\frac{\mu_{u_j}}{\mu}\right),$$

and

$$\frac{\mu_{u_i u_j}}{\mu} = \left(\frac{\mu_{u_j}}{\mu}\right)_{u_i} + \left(\frac{\mu_{u_i}}{\mu}\right)\left(\frac{\mu_{u_j}}{\mu}\right).$$

This fact would give rise to a lack of uniqueness, which must be avoided by observing the following rule: *For any particular coefficient of* $(\bar{a})$, *we must decide which of the two possible substitutions for* $\mu_{u_i u_j}/\mu$ *is to be used. Throughout the coefficient, that substitution must be used, the symbolic identities* (4) *being employed, if necessary.* For example, suppose that a certain coefficient of $(\bar{a})$ is of the form

$$\cdots + a\frac{\mu_{u_1 u_2}}{\mu} + \cdots + b\frac{\mu_{u_1 u_1 u_2}}{\mu} + c\frac{\mu_{u_1 u_2 u_2}}{\mu} + \cdots .$$

If $\mu_{u_1}/\mu = a'$, $\mu_{u_2}/\mu = b'$, then $\mu_{u_1 u_2}/\mu = a'_{u_2} + a'\, b'$, or $\mu_{u_1 u_2}/\mu = b'_{u_1} + a'\, b'$. Either substitution for $\mu_{u_1 u_2}/\mu$ may be chosen, but when chosen, must be used for $\mu_{u_1 u_1 u_2}/\mu$, $\mu_{u_1 u_2 u_2}/\mu$, and all other μ-fractions in the coefficient whose numerators are partial derivatives of $\mu_{u_1 u_2}$. That is, supposing that $\mu_{u_1 u_2}/\mu$ is chosen equal to $a'_{u_2} + a'\, b'$, b'_{u_1} must be replaced by a'_{u_2}, whenever b'_{u_1} or one of its derivatives appears in the coefficient.

Having observed this precaution, it is evident that we have formally obtained a unique form of the completely integrable system (a). We shall call this form a *pseudo-semi-canonical form* (A), of (a), and shall use capital letters for its coefficients. It is completely characterized by the n equations $B_i = 0$, $(i = 1, \cdots , n)$. The seminvariance of the coefficients of (A) may be established by the following argument.

Let us refer to the set of coefficients of the system (a) as $[a]$. Under the transformation $(2a)$ they are replaced by the set $[a_\mu]$, the first class coefficients of which are of the type (3). $[a]$ and $[a_\mu]$ may be regarded as the coefficients of any two systems of type (a) which are equivalent under the transformation $(2a)$. Apply to (a) and (a_μ) the transformation $y = \nu\bar{y}$, which is of the type $(2a)$. The coefficients $[a_\nu]$ and $[a_{\mu\nu}]$ which result have the following relations: (i) $[a_{\mu\nu}]$ are the same functions of $[a_\mu]$ and ν_{u_i}/ν, $(i = 1, \cdots, n)$, as $[a_\nu]$ are of $[a]$ and ν_{u_i}/ν; (ii) $[a_{\mu\nu}]$ are the same functions of $[a]$ and $(\mu\nu)_{u_i}/\mu\nu$, $(i = 1, \cdots, n)$, as $[a_\nu]$ are of $[a]$ and ν_{u_i}/ν.

In $[a_\nu]$ make n coefficients of the first class (see (3)) vanish by a suitable choice of ν_{u_i}/ν, $(i = 1, \cdots, n)$, as functions of $[a]$, regarding these functions (denoted by (ν_{u_i}/ν)), as independent unknowns. If we substitute (ν_{u_i}/ν) for ν_{u_i}/ν throughout $[a_\nu]$, and observe the above mentioned precaution concerning cross-derivatives, we obtain unique functions $[A]$ of the original coefficients $[a]$. The functions $[A]$ are also unique as a set.

Make the corresponding first class coefficients of $[a_{\mu\nu}]$ vanish. This will be accomplished by taking ν_{u_i}/ν equal to the same functions, $(\nu_{u_i}/\nu)'$, of $[a_\mu]$ as the (ν_{u_i}/ν) are of $[a]$. But this choice of the ν_{u_i}/ν is equivalent to taking $(\mu\nu)_{u_i}/\mu\nu$ equal to the same functions, $((\mu\nu)_{u_i}/\mu\nu)$, of $[a]$, so that the functions $((\mu\nu)_{u_i}/\mu\nu)$ and (ν_{u_i}/ν) are identical. The substitution of $(\nu_{u_i}/\nu)'$ for ν_{u_i}/ν throughout $[a_{\mu\nu}]$ will yield a set of unique functions, $[A]'$, of the coefficients $[a]$.

By virtue of the relation (i), the functions $[A]'$ are the same functions of $[a_\mu]$ and $(\nu_{u_i}/\nu)'$ as $[A]$ are of $[a]$ and (ν_{u_i}/ν). But $(\nu_{u_i}/\nu)'$ are the same functions of the coefficients $[a_\mu]$ as (ν_{u_i}/ν) are of the original coefficients $[a]$. Hence the functions $[A]'$ are the same functions of $[a_\mu]$ as $[A]$ are of $[a]$, where $[a_\mu]$ and $[a]$ are the coefficients of any two systems of the type (a) equivalent under the transformation $(2a)$.

In view of the relation (ii), the functions $[A]'$ are the same functions of $[a]$ and $((\mu\nu)_{u_i}/\mu\nu)$ as $[A]$ are of $[a]$ and (ν_{u_i}/ν). Since $((\mu\nu)_{u_i}/\mu\nu)$ and (ν_{u_i}/ν) are identical, we see that the functions $[A]'$ and $[A]$ are identical. Hence they are seminvariants.

We note that the number of seminvariant coefficients of a pseudo-semicanonical form is $pq - n$, exactly the number of seminvariants in a fundamental set as obtained by Green.* They are moreover independent, no two arising from the same coefficient of (a). Hence *the coefficients* $[A]$ *are a fundamental set of seminvariants.*

3. Semi-covariants

The transformation $(1a)$ leads us to two important sets of equations. One set is that by means of which the new coefficients (those of $(\bar{a})$) are expressed

* *Linear Dependence of Functions*, etc., l. c., section 6.

in terms of the old ones (those of (a)) and of the transformation function μ. The other is the set of equations (2), which express the new variables in terms of the old ones and of μ. There is a parity existing between these two sets which, apparently, has not usually been recognized.

To exhibit one aspect of this parity, let us make in equations (2) the substitutions $\mu_{u_i}/\mu = (\mu_{u_i}/\mu)$, $(i = 1, \cdots, n)$, which are determined by the equations $\bar{\beta}_i = 0$, $(i = 1, \cdots, n)$. In these substitutions, we disregard the factor μ. *The functions*

$$Y = y, \qquad Y_{u_i} = y_{u_i} + (\mu_{u_i}/\mu)\, y, \qquad (i = 1, \cdots, n), \tag{5}$$

are relative covariants. We shall prove this proposition by an argument parallel to that made in section 2 for the seminvariance of the functions $[A]$.

When the transformation (2*a*) is made, the set of variables

$$[y]: \quad y,\ y_{u_i},\ \text{etc.} \qquad (i = 1, \cdots, n),$$

is replaced by the set $[y_\mu]$, given by equations (2). When we apply the transformation $\bar{y} = \nu y$ to the systems (a) and (a_μ), we get the new sets of variables $[y_\nu]$ and $[y_{\mu\nu}]$, respectively. If we substitute $\nu_{u_i}/\nu = (\nu_{u_i}/\nu)$ in $[y_\nu]$, there results (if we disregard the factor ν) the unique set of functions $[Y]$. Similarly, the substitution of $\nu_{u_i}/\nu = (\nu_{u_i}/\nu)'$ in $[y_{\mu\nu}]$ gives rise to the set of functions $[Y]'$. As in section 2, we can prove the following statements, which suffice to establish the truth of the proposition: (i) The set $[Y]'$ are the same functions (neglecting a factor) of $[y_\mu]$ and $[a_\mu]$ as $[Y]$ are of $[y]$ and $[a]$; (ii) Except for a factor, the set $[Y]'$ are the same functions of $[y]$ and $((\mu\nu)_{u_i}/\mu\nu)$ as $[Y]$ are of $[y]$ and (ν_{u_i}/ν). Hence the functions $[Y]'$ and $[Y]$ are identical, except for a factor.

An induction proof may be made as follows: We first show that

$$\overline{\left(\frac{\mu_{u_i}}{\mu}\right)} = \left(\frac{\mu_{u_i}}{\mu}\right) - \frac{\mu_{u_i}}{\mu} \qquad (i = 1, \cdots, n).$$

In order to do this, let us refer to the definitions of (μ_{u_i}/μ). Since these functions arise from n equations of the type (3), we have

$$B_g = \beta_g + \sum_{i=1}^{n} p_{gi}\, \alpha_{gi} \left(\frac{\mu_{u_i}}{\mu}\right) = 0 \qquad (g = 1, \cdots, n). \tag{6}$$

But, since the combinations B_g are seminvariants,

$$\bar{\beta}_g + \sum_{i=1}^{n} p_{gi}\, \alpha_{gi} \overline{\left(\frac{\mu_{u_i}}{\mu}\right)} = 0 \qquad (g = 1, \cdots, n),$$

the $p_{gi}\, \alpha_{gi}$ being also seminvariants. By use of (3) and (6), we may put these equations in the form

$$\sum_{i=1}^{n} p_{gi}\, \alpha_{gi} \left[\overline{\left(\frac{\mu_{u_i}}{\mu}\right)} - \left(\frac{\mu_{v_i}}{\mu}\right) + \frac{\mu_{u_i}}{\mu} \right] = 0 \qquad (g = 1, \cdots, n).$$

The determinant of this system does not vanish, since we have assumed that the equations (6) are solvable for the (μ_{u_i}/μ). Therefore,

$$\left(\overline{\frac{\mu_{u_i}}{\mu}}\right) = \left(\frac{\mu_{u_i}}{\mu}\right) - \frac{\mu_{u_i}}{\mu} \qquad (i = 1, \cdots, n).$$

As a result of this lemma, we have

$$\overline{Y}_{u_i} = \bar{y}_{u_i} + \left(\overline{\frac{\mu_{u_i}}{\mu}}\right)\bar{y} = \mu\left[y_{u_i} + \frac{\mu_{u_i}}{\mu}y\right] + \left[\left(\frac{\mu_{u_i}}{\mu}\right) - \frac{\mu_{u_i}}{\mu}\right]\mu y = \mu Y_{u_i}$$
$$(i = 1, \cdots, n).$$

It is also evident from (2) that $\overline{Y} = \mu Y$. Now assume that

(7) $$\overline{Y}_{l_1, \cdots, l_n} = \mu Y_{l_1, \cdots, l_n}.$$

From the law of formation of the functions (5), it follows that

$$\begin{aligned}
&Y_{l_1, \cdots, l_{i-1}, l_i+1, l_{i+1}, \cdots, l_n} \\
&= \sum \binom{l_1}{p_1} \cdots \binom{l_n}{p_n} \Bigg\{ \left(\frac{\mu_{l_1, \cdots, l_{i-1}, l_i+1, l_{i+1}, \cdots, l_n}}{\mu}\right) y_{l_1-p_1, \cdots, l_n-p_n} \\
&\qquad\qquad + \left(\frac{\mu_{p_1, \cdots, p_n}}{\mu}\right) y_{l_1-p_1, \cdots, l_{i-1}-p_{i-1}, l_i-p_i+1, l_{i+1}-p_{i+1}, \cdots, l_n-p_n} \Bigg\} \\
&= \sum \binom{l_1}{p_1} \cdots \binom{l_n}{p_n} \Bigg\{ \Bigg[\left(\frac{\mu_{l_1, \cdots, l_n}}{\mu}\right)_{u_i} \\
&\qquad\qquad + \left(\frac{\mu_{u_i}}{\mu}\right)\left(\frac{\mu_{l_1, \cdots, l_n}}{\mu}\right)\Bigg] y_{l_1-p_1, \cdots, l_n-p_n} \\
&\qquad\qquad + \left(\frac{\mu_{p_1, \cdots, p_n}}{\mu}\right) y_{l_1-p_1, \cdots, l_{i-1}-p_{i-1}, l_i-p_i+1, l_{i+1}-p_{i+1}, \cdots, l_n-p_n} \Bigg\} \\
&= \frac{\partial}{\partial u_i} \sum \binom{l_1}{p_1} \cdots \binom{l_n}{p_n} \left(\frac{\mu_{l_1, \cdots, l_n}}{\mu}\right) y_{l_1-p_1, \cdots, l_n-p_n} \\
&\qquad\qquad + \left(\frac{\mu_{u_i}}{\mu}\right) \sum \binom{l_1}{p_1} \cdots \binom{l_n}{p_n} \left(\frac{\mu_{l_1, \cdots, l_n}}{\mu}\right) y_{l_1-p_1, \cdots, l_n-p_n} \\
&= \frac{\partial}{\partial u_i} Y_{l_1, \cdots, l_n} + \left(\frac{\mu_{u_i}}{\mu}\right) Y_{l_1, \cdots, l_n},
\end{aligned}$$

where all summations are extended over $p_1, p_2, \cdots, p_n$; $p_1 = 0, \cdots, l_1$; $\cdots$; $p_n = 0, \cdots, l_n$. Hence, by assumption (7),

$$\overline{Y_{l_1, \cdots, l_{i-1}, l_i+1, l_{i+1}, \cdots, l_n}} = \frac{\partial}{\partial u_i}\overline{Y_{l_i, \cdots, l_n}} + \left(\overline{\frac{\mu_{u_i}}{\mu}}\right)\overline{Y_{l_i, \cdots, l_n}}$$

$$= \frac{\partial}{\partial u_i}(\mu Y_{l_1, \cdots, l_n}) + \left[\left(\frac{\mu_{u_i}}{\mu}\right) - \frac{\mu_{u_i}}{\mu}\right]\mu Y_{l_1, \cdots, l_n}$$

$$= \mu \left\{ \frac{\partial}{\partial u_i} Y_{l_1, \cdots, l_n} + \left(\frac{\mu_{u_i}}{\mu}\right) Y_{l_1, \cdots, l_n} \right\}$$

$$= \mu Y_{l_1, \cdots, l_{i-1}, l_i+1, l_{i+1}, \cdots, l_n} \qquad (i = 1, \cdots, n),$$

which completes the induction proof.

The form of (5) shows that all the functions there listed would be independent except for the system (a). By means of (i) all y-derivatives are expressed linearly in terms of the primary derivatives. But (5) may be solved for y, y_{u_i}, etc., in terms of Y, Y_{u_i}, etc. Hence all semi-covariants (5) are functions of those which correspond to the primary derivatives, and of seminvariants.

A uniquely determined set of functions may also be obtained from the coefficients of $(\bar{a})$ by the substitutions

$$\mu_{u_i}/\mu = - y_{u_i}/y \qquad (i = 1, \cdots, n),$$

which are suggested by equations (2). This set of functions will be characterized by the relations

$$Y_{u_i} = 0 \qquad (i = 1, \cdots, n).$$

As a result of the uniqueness of this set, they are semi-covariants.

4. Application to algebraic semi-covariants

We may regard a pseudo-semi-canonical form of a completely integrable system of linear homogeneous partial differential equations as a formal analogue of the reduced binary form, and of the semi-canonical form of the linear homogeneous ordinary differential equation. Indeed, the suggestion contained in the last paragraph of section 3 may be carried out in the theory of binary forms.

When the binary p-ic*

$$f = a_0 x^p + p a_1 x^{p-1} y + \cdots + a_p y^p$$

is subjected to the transformation

$$x = \xi + n\eta, \qquad y = \eta,$$

the variables, $x^p, x^{p-1} y, \cdots, y^p$, are changed in accordance with the equations

$$(8) \qquad x^p = (\xi + n\eta)^p, \qquad x^{p-1} y = (\xi + n\eta)^{p-1}\eta, \qquad \cdots, \qquad y^p = \eta^p,$$

while the new coefficients are expressed in terms of the old by means of the equations

$$(9) \qquad a_0' = a_0, \qquad a_1' = a_1 + n a_0, \qquad a_2' = a_2 + 2n a_1 + n^2 a_0, \qquad \text{etc.}$$

* Cf. Dickson, *Algebraic Invariants*, 1914, p. 47.

The substitution in (9) of $n = -\xi/\eta$, suggested by (8), will yield a set of independent semi-covariants.

5. Effect of the transformation of the independent variables upon the coefficients of (a)

Let the independent variables be transformed by means of the equations

$$\hat{u}_i = U_i(u_i) \qquad (i = 1, \cdots, n). \tag{10}$$

The effect upon the variables y, y_{u_i}, etc., is shown by the following equations:

$$\begin{aligned}
&y = \hat{y}, \qquad y_{u_i} = U'_i\,\hat{y}_{u_i}, \qquad y_{u_i u_i} = U'^2_i\,\hat{y}_{u_i u_i} + U''_i\;\hat{y}_{u_i},\\
&y_{u_i u_j} = U'_i\,U'_j\,\hat{y}_{u_i u_j}, \qquad (i \neq j),\\
&y_{u_i u_i u_i} = U'^3_i\,\hat{y}_{u_i u_i u_i} + 3U'_i\,U''_i\,\hat{y}_{u_i u_i} + U'''_i\,\hat{y}_{u_i},\\
&y_{u_i u_i u_j} = U'^2_i\,U'_j\,\hat{y}_{u_i u_i u_j} + U''_i\,U'_j\,\hat{y}_{u_i u_j}, \qquad (i \neq j),\\
&y_{u_i u_j u_k} = U'_i\,U'_j\,U'_k\,\hat{y}_{u_i u_j u_k}, \qquad (i \neq j \neq k), \text{ etc.},\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&y_{l_1, \cdots, l_n} = \hat{y}_{l_1, \cdots, l_n} \cdot U'^{l_1}_1 \cdots U'^{l_n}_n\\
&+ \sum_{g=1}^{n} q_g\,\hat{y}_{l_1, \cdots, l_{g-1}, l_g - 1, l_{g+1}, \cdots, l_n} \cdot U'^{l_1}_1 \cdots U'^{l_{g-1}}_{g-1}\,U'^{l_g - 2}_g U'^{l_{g+1}}_{g+1}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdots U'^{l_n}_n\,U''_g + \cdots,
\end{aligned} \tag{11}$$

where $\hat{y}_{u_i} = \partial\hat{y}/\partial\hat{u}_i$, etc., and the q_g are integers (including zero). An easy induction suffices to show the correctness of the last expression of (11).

It is obvious from (11) that when the expressions there given for y, y_{u_i}, etc., are substituted in (a), and the resulting equations collected in a system $(\hat{a})$ of the same type, a given y-derivative may give increments only to those terms of its equation whose y-derivatives yield the given y-derivative by differentiation. For example, the new coefficient of $\hat{y}_{u_i u_j}$ in any equation of $(\hat{a})$ will be due to the old $y_{u_i u_j}$, $y_{u_i u_i u_j}$, etc., provided these derivatives are present in this equation. In this respect, the effects of the transformations $(1a)$ and (10) are similar. There is this difference to be noticed, however. The right member of any equation of (2) will contain *all* y-derivatives which yield by differentiation the derivative in the left member of this equation. Of equations (11), on the other hand, there is only one, namely the first, of which this is true. For example, in the expression for $y_{u_i u_j}$, $(i \neq j)$, there occurs none of the variables $\hat{y}$, $\hat{y}_{u_i}$, $\hat{y}_{u_j}$, all of which would appear in the corresponding equation of (2).

In order to show more explicitly the effect of the transformation (10) upon (a), let one of the equations of (a) be the following:*

$$
\tag{12}
\begin{aligned}
y_{l_1+1,\, l_2,\, \cdots,\, l_n} &= \sum_{g=2}^{n} \alpha_g\, y_{l_1,\, \cdots,\, l_{g-1},\, l_g+1,\, l_{g+1},\, \cdots,\, l_n} + \beta y_{l_1,\, \cdots,\, l_n} + \cdots \\
&\quad + \sum_{g=1}^{n} \epsilon_g\, y_{m_1 \cdots\, m_{g-1},\, m_g+1,\, m_{g+1},\, \cdots,\, m_n} + \zeta y_{m_1,\, \cdots,\, m_n} + \cdots,
\end{aligned}
$$

where $\Sigma l_i > \Sigma m_i$; and where the β-term is typical of all first class terms, while the ζ-term typifies all terms of higher class whose coefficients will, under the transformation (10), receive additional terms only from the y-derivatives

$$
y_{m_1,\, \cdots,\, m_{g-1}\, m_g+1,\, m_{g+1},\, \cdots,\, m_n} \qquad (g = 1, \cdots, n),
$$

of order one higher than $y_{m_1,\, \cdots,\, m_n}$.

The last equation of (11) gives us, as the expression for the y-derivatives from the ϵ-terms:

$$
\tag{13}
\begin{aligned}
&y_{m_1,\, \cdots,\, m_{g-1},\, m_g+1,\, m_{g+1},\, \cdots,\, m_n} \\
&\quad = \hat{y}_{m_1,\, \cdots,\, m_{g-1},\, m_g+1,\, m_{g+1},\, \cdots,\, m_n} \cdot U_1'^{\,m_1} \cdots U_{g-1}'^{\,m_{g-1}}\, U_g'^{\,m_g+1}\, U_{g+1}'^{\,m_{g+1}} \cdots U_n'^{\,m_n} \\
&\qquad + q\hat{y}_{m_1,\, \cdots,\, m_n} \cdot U_1'^{\,m_1} \cdots U_{g-1}'^{\,m_{g-1}}\, U_g'^{\,m_g-1}\, U_{g+1}'^{\,m_{g+1}} \cdots U_n'^{\,m_n}\, U_g'' + \cdots,
\end{aligned}
$$

where we are assuming q, which cannot be negative, to be different from zero.

Suppose that one of the ϵ_g-terms receives an increment, by the transformation (10), from a y-derivative of higher order than that in the ϵ_g-term, say from $y_{l_1,\, \cdots,\, l_n}$. The expression for $y_{l_1,\, \cdots,\, l_n}$ will be found by successive differentiation of (13). But since the second term of the right member of (13) is a product, repeated differentiation will not make the $y_{m_1,\, \cdots,\, m_n}$ term disappear unless at least one of the U_i' and its derivatives are missing from the factor

$$
U_1'^{\,m_1} \cdots U_{g-1}'^{\,m_{g-1}}\, U_g'^{\,m_g-1}\, U_{g+1}'^{\,m_{g+1}} \cdots U_n'^{\,m_n}\, U_g''.
$$

Equations (11) show that this would mean that U_i' would also be absent from the coefficient of every y-derivative in the complete expression (13). Therefore, in forming out of (13) the expression for $y_{l_1,\, \cdots,\, l_n}$, not only $y_{m_1,\, \cdots,\, m_n}$, but $y_{m_1,\, \cdots,\, m_{g-1},\, m_g+1,\, m_{g+1},\, \cdots,\, m_n}$ as well, disappears. Hence, if an ϵ_g-term receives an increment from a higher order derivative in its equation, the ζ-term will likewise receive an increment from the same higher order derivative. But this is contrary to the assumption that the ζ-term receives increments only from the derivatives in the ϵ_g-terms. Therefore the ϵ_g are unaltered, except for a factor, by the transformation (10).

When the equations (11) are substituted in (12), we obtain the equation

$$
\hat{y}_{l_1+1,\, l_2,\, \cdots,\, l_n} \cdot U_1'^{\,l_1+1}\, U_2'^{\,l_2} \cdots U_n'^{\,l_n} + q_1\, \hat{y}_{l_1,\, \cdots,\, l_n} \cdot U_1'^{\,l_1-1}\, U_2'^{\,l_2} \cdots U_n'^{\,l_n}\, U_1'' + \cdots
$$

* It is not material to the argument which one of the derivatives of order $1 + \Sigma l_i$ is the left member.

$$
\begin{aligned}
&= \sum_{g=2}^{n} \alpha_g \,(\hat{y}_{l_1, \cdots, l_{g-1}, l_g+1, l_{g+1}, \cdots, l_n} \cdot U_1'^{\,l_1} \cdots U_{g-1}'^{\,l_{g-1}} U_g'^{\,l_g+1} U_{g+1}'^{\,l_{g+1}} \cdots U_n'^{\,l_n} \\
&\quad + q_g \,\hat{y}_{l_1, \cdots, l_n} \cdot U_1'^{\,l_1} \cdots U_{g-1}'^{\,l_{g-1}} U_g'^{\,l_g-1} U_{g+1}'^{\,l_{g+1}} \cdots U_n'^{\,l_n} U_g'') + \cdots \\
&\quad + \beta \hat{y}_{l_1, \cdots, l_n} \cdot U_1'^{\,l_1} \cdots U_n'^{\,l_n} + \cdots \\
&\quad + \sum_{g=1}^{n} \epsilon_g \,(\hat{y}_{m_1, \cdots, m_{g-1}, m_g+1, m_{g+1}, \cdots, m_n} \cdot U_1'^{\,m_1} \cdots U_{g-1}'^{\,m_{g-1}} U_g'^{\,m_g+1} U_{g+1}'^{\,m_{g+1}} \cdots U_n'^{\,m_n} \\
&\quad + q_g' \,\hat{y}_{m_1, \cdots, m_n} \cdot U_1'^{\,m_1} \cdots U_{g-1}'^{\,m_{g-1}} U_g'^{\,m_g-1} U_{g+1}'^{\,m_{g+1}} \cdots U_n'^{\,m_n} U_g'') + \cdots \\
&\quad + \zeta \hat{y}_{m_1, \cdots, m_n} \cdot U_1'^{\,m_1} \cdots U_n'^{\,m_n} + \cdots .
\end{aligned}
$$

Upon dividing by $U_1'^{\,l_1+1} U_2'^{\,l_2} \cdots U_n'^{\,l_n}$, and transposing, we see that

$$
\begin{gathered}
\hat{\alpha}_g = \frac{U_g'}{U_1'} \alpha_g , \qquad \hat{\beta} = \frac{1}{U_1'} \left(\beta - q_1 \eta_1 + \sum_{g=2}^{n} q_g \alpha_g \eta_g \right), \\
\hat{\epsilon}_g = \frac{U_1'^{\,m_1} \cdots U_{g-1}'^{\,m_{g-1}} U_g'^{\,m_g+1} U_{g+1}'^{\,m_{g+1}} \cdots U_n'^{\,m_n}}{U_1'^{\,l_1+1} U_2'^{\,l_2} \cdots U_n'^{\,l_n}} \cdot \epsilon_g , \\
\hat{\zeta} = \frac{U_1'^{\,m_1} \cdots U_n'^{\,m_n}}{U_1'^{\,l_1+1} U_2'^{\,l_2} \cdots U_n'^{\,l_n}} \left(\zeta + \sum_{g=1}^{n} q_g' \epsilon_g \eta_g \right),
\end{gathered} \tag{14}
$$

where

$$\eta_g \equiv U_g'' / U_g' .$$

All zeroth class coefficients, therefore, are unaltered, except for a factor, by the transformation (10). Since they are known to be unchanged by the transformation (1a), they are relative invariants.

If we write the $\hat{\beta}$ expression as

$$\hat{\beta} = \frac{1}{U_1'} \left(\beta + \sum_{g=1}^{n} q_g \alpha_g \eta_g \right) \qquad (\alpha_1 \equiv -1), \tag{15}$$

we may take this as typical (except for the factor), not only of all first class coefficients of $(\hat{a})$, but also of all coefficients, of whatever class, which are altered (factors disregarded) only by linear combinations of η_i $(i = 1, \cdots, n)$. In all such cases, the coefficients in these linear combinations are unchanged, except for a factor, by the transformation (10).

Consider any pseudo-semi-canonical form (A) of (a), characterized by the relations $B_g = 0$ $(g = 1, \cdots, n)$. We wish to observe the effect of (10) upon its coefficients. This may be done by direct substitution, but it is desirable for our purpose to indicate a different method.

The coefficients, $\bar{\beta}_g$, of $(\bar{a})$, whose vanishing characterizes the pseudo-semi-canonical form, may or may not be unaltered by the transformation (10). In the former case, the pseudo-semi-canonical form will be preserved, and equations (14), which describe the effect of (10) upon the original coefficients of (a), will serve also to indicate its effect upon the coefficients of the pseudo-semi-canonical form (A). We need only read A for α, B for β, etc.

In the latter case, the pseudo-semi-canonical form is violated, and must be restored by a second application of the transformation (1a). Equation (15) shows that the undesired increments, which must be removed from the $\hat{B}_g$, are at worst linear combinations of $\eta_1, \cdots, \eta_n$ with relative invariant multipliers. In this second application of (1a) to the form $(\hat{A})$, $\mu_{u_1}/\mu, \cdots, \mu_{u_n}/\mu$, must be replaced by linear combinations of the kind just described.

We recall that the coefficients of $(\bar{a})$ are equal to the corresponding coefficients of (a) plus linear combinations of μ-fractions with multipliers taken from among the original coefficients of (a). Hence, since we are applying (1a) to a form of (a) whose coefficients are seminvariants, we see that the coefficients of $(\hat{A})$ must, certain factors neglected, equal the corresponding coefficients of (A) plus functions of seminvariants and of the $\eta_1, \cdots, \eta_n$ and their derivatives.

The foregoing paragraphs throw some light on a question which arose in the author's previous paper.* Certain pseudo-semi-canonical forms in the special cases there discussed, not only yielded simpler seminvariants than did the classical process, but also produced a relatively large number of seminvariants which were relative invariants as well.

Equations (11) show that the coefficients of the primary derivative y will be unaltered by the transformation (10), except for a factor. For this reason, the system $(\hat{a})$ is usually much simpler than $(\hat{A})$. However, when a pseudo-semi-canonical form (A) has been used which is not violated by the transformation (10), the systems $(\hat{A})$ and $(\hat{a})$ are identical, and we may be sure that at least the coefficients of y in (A) will be relative invariants.

6. Pseudo-canonical forms and invariants

Equation (15), factors neglected, is typical of all coefficients of $(\hat{A})$ which have as increments only linear combinations of $\eta_1, \cdots, \eta_n$. Let us choose† n such coefficients and form the equations $\hat{B}_g = 0$, $(g = 1, \cdots, n)$. Assuming that the B_g have been so chosen as to permit, we solve the equations for $\eta_1, \cdots, \eta_n$, regarding these as independent unknowns. When these solutions, denoted by $(\eta_1), \cdots, (\eta_n)$, are substituted for the η's throughout $(\hat{A})$, we obtain (formally) a form $(\mathfrak{A})$ of (A), whose coefficients are unique except for certain factors $U_1'^{m_1} \cdots U_n'^{m_n}$. We shall call this form, which is characterized by the relations $\mathfrak{B}_g = 0$ $(g = 1, \cdots, n)$, a *pseudo-canonical form* of (a), and denote its coefficients by German capitals. By an argument

* A. L. Nelson, l. c. Cf. section 6.

† Cf. the discussion in section 8.

analogous to that used in section 2, we are able to prove that these coefficients are unchanged, factors neglected, by the transformation (10). Since, in addition, the coefficients of $(\hat{A})$ are functions of seminvariants and of $\eta_1, \cdots, \eta_n$, and the $(\eta_1), \cdots, (\eta_n)$ are seminvariants, it follows that *the coefficients of* $(\mathfrak{A})$ *are relative invariants.*

The invariants thus formed are in number $2n$ less than the total number of original coefficients of (a). That is to say, our set is n short of the number required for a fundamental set. Those we have are evidently independent, and it remains to indicate how n more may be formed, independent of each other and of those already obtained. These supplementary invariants may be found by a device exactly analogous to that employed by Wilczynski.*

In section 2 it was proved that

$$\left(\overline{\frac{\mu_{u_i}}{\mu}}\right) = \left(\frac{\mu_{u_i}}{\mu}\right) - \frac{\mu_{u_i}}{\mu}.$$

In an exactly similar manner, it may be shown that, except for a factor,

$$(\hat{\eta}_i) = (\eta_i) - \eta_i \qquad (i = 1, \cdots, n). \tag{16}$$

The second equation of (14) shows that $\alpha_g(\eta_g)$ must be transformed by (10) in accordance with the equations

$$\hat{\alpha}_g(\hat{\eta}_g) = \frac{1}{U_1'}[\alpha_g(\eta_g) + \cdots] \qquad (g = 1, \cdots, n).$$

Substitution of the first equation of (14) gives

$$(\hat{\eta}_g) = \frac{1}{U_g'}[(\eta_g) + \cdots], \tag{17}$$

for all (η_g) which arise from first class coefficients. Similarly, the third and fourth equations of (14) prove that for all other (η_g), the same factor, $1/U_g'$, occurs. Hence, combining (16) and (17), we see that the complete expression for $(\hat{\eta}_i)$ is

$$(\hat{\eta}_i) = \frac{1}{U_1'}[(\eta_i) - \eta_i] \qquad (i = 1, \cdots, n).$$

Choose any other relative invariant, θ, where

$$\hat{\theta} = U_1'^{l_1} \cdots U_n'^{l_n} \theta.$$

By differentiation, we obtain

$$\hat{\theta}_{u_i} = \frac{1}{U_1'} \frac{\partial}{\partial u_i}[U_1'^{l_1} \cdots U_n'^{l_n} \theta]$$

* *One-parameter Families*, etc., l. c. Cf. equations (25).

$$= \frac{1}{U_i'}[U_1'^{l_1} \cdots U_n'^{l_n} \theta_{u_i} + l_i U_1'^{l_1} \cdots U_{i-1}'^{l_{i-1}} U_i'^{l_i-1} U_{i+1}'^{l_{i+1}} \cdots U_n'^{l_n} U_i'' \theta] \qquad (i = 1, \cdots, n),$$

so that

$$\frac{\hat{\theta}_{u_i}}{\hat{\theta}} = \frac{1}{U_1'}\left[\frac{\theta_{u_i}}{\theta} + l_i \eta_i\right] \qquad (i = 1, \cdots, n).$$

The functions

$$l_i(\eta_i) + \theta_{u_i}/\theta \qquad (i = 1, \cdots, n),$$

will be n new relative invariants, evidently independent of each other. They are also independent of the invariants which are coefficients of the pseudo-canonical form, because none of these coefficients possessed any of the (η_i) as leaders.

We recall here general principles underlying the formation of the coefficients of the pseudo-canonical form $(\mathfrak{A})$. Neglecting factors, the zeroth class coefficients of $(\mathfrak{A})$ are equal to the corresponding coefficients of the pseudo-semi-canonical form (A) from which $(\mathfrak{A})$ was formed. The non-vanishing coefficients of the first class are linear combinations of first class seminvariant coefficients of (A), the coefficients in these linear combinations being numerical multiples of zeroth class invariants. Each of the invariants of any particular class higher than the first, is equal to the corresponding seminvariant coefficient of (A) plus combinations of lower class seminvariant coefficients.

Hence, in the set of invariants, including the n invariants

$$(18) \qquad l_i(\eta_i) + \theta_{u_i}/\theta \qquad (i = 1, \cdots, n),$$

we may solve for the corresponding seminvariant leaders. As a result, any invariant whatever of the original completely integrable system (a), being known to be a function of the fundamental seminvariants and their derivatives, is seen to be a function of the invariant coefficients of $(\mathfrak{A})$ and of the invariants (18). Therefore *these invariants are a fundamental set.*

7. Covariants

The semi-covariants (5) may be regarded as the variables of the pseudo-semi-canonical form (A). Under the transformation (10), they will be altered in accordance with equations (11), reading Y, Y_{u_i}, etc., for y, y_{u_i}, etc. We may solve these equations, so as to express $\hat{Y}$, $\hat{Y}_{u_i}$, etc., in terms of Y, Y_{u_i}, etc.

However, if the pseudo-semi-canonical form is violated by the transformation (10), it must be restored by a transformation of the type (1a). In other words, equations (2), with $\mu_{u_1}/\mu, \cdots, \mu_{u_n}/\mu$ so chosen as to put the system $(\hat{A})$ in the pseudo-semi-canonical form again, must in such cases follow the transformation (10), in order to give in final form the effect of (10) upon the semi-covariants (5). This will give us expressions for $\hat{Y}$, $\hat{Y}_{u_i}$, etc., as functions of the semi-covariants (5), the seminvariant coefficients of (A), and of $U'_1, \cdots, U'_n, \eta_1, \cdots, \eta_n$.

The substitutions $\eta_i = (\eta_i)$ $(i = 1, \cdots, n)$, which produce the pseudo-canonical form $(\mathfrak{A})$ of (A), will also give a unique set of equations which express $\hat{Y}$, $\hat{Y}_{u_i}$, etc., in terms of Y, Y_{u_i}, etc. Neglecting the factors $U_1'^{l_1}, \cdots, U_n'^{l_n}$, let us call these expressions $\mathfrak{Y}$, $\mathfrak{Y}_{u_i}$, etc. *They are relative covariants*, and are characterized as a set by the relations $\mathfrak{B}_g = 0$ $(g = 1, \cdots, n)$, which are characteristic of the pseudo-canonical form $(\mathfrak{A})$.

Those of the $\mathfrak{Y}$, $\mathfrak{Y}_{u_i}$, etc., up to a certain order, which is determined by the left members of the completely integrable system (a), are evidently independent. *All covariants are functions of these independent ones and of invariants.*

It is sufficient merely to remark that fundamental sets of simultaneous invariants and seminvariants of sets of completely integrable systems of partial differential equations may be constructed by the method outlined in the preceding sections. Examples of such sets, of which the simultaneous invariants might be of great interest, are suggested by two theorems of Koenigs* concerning the perspective plane nets of the asymptotic lines and of conjugate systems of curves on curved surfaces. Further examples are furnished by the Laplace suite, as applied to plane nets,† to conjugate nets of curves on surfaces,‡ and to congruences of straight lines.§

8. Concerning the possibility of securing a pseudo-canonical form

It must be emphasized that the "reduction" of a completely integrable system (a) to a pseudo-semi-canonical form is possible only when there are present n first class coefficients of $(\bar{a})$ of the type (3). Moreover, even after such a pseudo-semi-canonical form (A) has been obtained, the further "reduction" of (A) to a pseudo-canonical form can be accomplished only when we have n additional coefficients of $(\hat{A})$, of the first and higher classes, of the type (15), not corresponding to the n coefficients of $(\bar{a})$ previously used.

It would be very desirable, therefore, to show that the situation just described always arises. Unfortunately, this very general conclusion cannot be established. On the contrary, an important example is at hand to prove the impossibility of such a conclusion. We are, however, able to show that for a large class of completely integrable systems, the method with which this paper has concerned itself, may be used.

* Paris Comptes Rendus, vol. 114 (1892), pp. 55–57; p. 728.

† E. J. Wilczynski, *One-parameter Families*, etc., l. c., section 4. See also A. L. Nelson, *Quasi-periodicity of Asymptotic Plane Nets*, Bulletin of the American Mathematical Society, vol. 22 (1916), pp. 445–455.

‡ G. M. Green, *Projective Differential Geometry of One-parameter Families of Space Curves*, etc., l. c., section 6.

§ E. J. Wilczynski, *Sur la Théorie Générale des Congruences*, l. c.

Suppose that the highest order derivative occurring in the completely integrable system is of order k, and that all kth order derivatives of y are present. For the purpose of our discussion, it is immaterial whether all the kth order derivatives are the left members of the equations of the system, or some are primary derivatives. There are in all*

$$H_k^n = C_k^{n+k-1} = \frac{n(n+1)\cdots(n+k-1)}{k!}$$

of these partial derivatives of order k. Among the cross-derivatives of this order will be a certain number of the type

$$(19) \qquad y_{l_1, \cdots, l_{r-1}, 1, l_{r+1}, \cdots, l_{s-1}, 1, l_{s+1}, \cdots, l_{t-1}, 1, l_{t+1}, \cdots, l_n},$$

which have been obtained by one and but one differentiation with respect to certain of the independent variables, u_r, u_s, u_t, etc. Each such derivative must yield an increment to the coefficient of every derivative of lower order out of which (19) comes by differentiation. In particular, the coefficients of the first class derivatives

$$(20) \qquad \begin{gathered} y_{l_1, \cdots, l_{r-1}, 0, l_{r+1}, \cdots, l_n}, \; y_{l_1, \cdots, l_s\ 1, 0, l_{s+1}, \cdots, l_n}, \\ y_{l_1, \cdots, l_{t-1}, 0, l_{t+1}, \cdots, l_n}, \text{ etc.,} \end{gathered}$$

in the system $(\bar{a})$, will receive, respectively, the increments

$$(21) \qquad \mu_{u_r}/\mu, \; \mu_{u_s}/\mu, \; \mu_{u_t}/\mu, \text{ etc.}$$

Let us try to count the derivatives of the type (19), listing them under n heads, according as u_1, u_2, $\cdots$, or u_n, is the independent variable which is represented just once in (19). Except for duplications there would be $n \cdot H_{k-1}^{n-1}$ of these derivatives. However, when a certain derivative would be listed under a number of these heads, that single derivative will yield suitable increments to the same number of first class coefficients. Hence there are exactly $n \cdot H_{k-1}^{n-1}$ increments of type (21) to coefficients of first class derivatives (20) from k-th order derivatives (19).

For $n = 2$, this number is exactly 2, for all values of k. For $k = 2$, it is equal to $n(n-1)$. Moreover, H_{k-1}^{n-1} increases with either n or k. Therefore, for $n > 1$, $k > 1$, there will be at least n coefficients of the first class of $(\bar{a})$, of the required type, which will serve to determine a pseudo-semi-canonical form. Equations (1) show that such a pseudo-semi-canonical form will be undisturbed by the transformation (10).

It must be remembered that, in general, the $n \cdot H_{k-1}^{k-1}$ coefficients we have considered are by no means the only possible ones for our purpose. Derivatives of type (19) may yield suitable increments to other coefficients than those

* Chrystal, *Algebra*, Part II, sec. 10, p. 10.

of type (20). Moreover, kth order derivatives of other types than (19) may give rise to proper increments.

Having obtained a pseudo-semi-canonical form (A), determined as outlined above, our next concern is with the system $(\hat{A})$. But, as has been remarked in section 5, it will suffice to discuss $(\hat{a})$. Consider the n " straight " derivatives

$$y_{u_j^k} \quad (j = 1, \cdots, n),$$

of order k, in the system (a). Each of these will give an increment, $- q_j \eta_j$, to the coefficient of $y_{u_j^{k-1}}$ in its equation. None of these straight derivatives have been used in the determination of the pseudo-semi-canonical form. Hence they may be made use of to secure a pseudo-canonical form.

In addition to the coefficients of $(\hat{A})$ suggested, we have available, in general, certain suitably incremented coefficients of $(\hat{A})$ which come from cross-derivatives. Indeed, the cross-derivatives (19) yield increments of the proper type to certain first and higher class coefficients of (A) which have not been used in the determination of the pseudo-semi-canonical form.

The above discussion shows that one or more pseudo-canonical forms may surely be obtained in completely integrable systems of the rather regular type assumed, *provided all the kth order derivatives are present.* Indeed, especially for larger values of n or k, we are assured a large freedom in the choice of the $2n$ coefficients whose vanishing characterizes a pseudo-canonical form. This indicates that pseudo-canonical forms may be obtained in a great number of less regular completely integrable systems.

In cases where some of the kth order derivatives are primary derivatives, however, one or more of them may be removed by a preliminary transformation of the type

$$\hat{u}_i = \phi_i(u_1, \cdots, u_n) \qquad (i = 1, \cdots, n),$$

which refers the configuration to a particular set of parametric curves. If a sufficient number of these primary derivatives are removed from all of the equations of the completely integrable system, the number of suitably incremented coefficients of $(\bar{a})$ or $(\hat{A})$ may become less than n, so that the pseudo-semi-canonical form, or the pseudo-canonical form, or both, cannot be obtained.

An example of such systems is the one used by Wilczynski* in the study of curved surfaces referred to their asymptotic lines. It is only possible, in this case, to obtain a pseudo-semi-canonical form, which coincides with his canonical form.

THE UNIVERSITY OF MICHIGAN

* E. J. Wilczynski, *Projective Differential Geometry of Curved Surfaces,* l. c. Cf. equations (27), (37), (38) and (46).

ARITHMETICAL PARAPHRASES[*] (II)

BY

E. T. BELL

IV. The fundamental series for paraphrase

1. A former paper under the same title is continued.[†] We first collect a few Fourier developments in a specially prepared form immediately suitable for paraphrase, and then apply the theory of the preceding paper to read off from them a few of their innumerable arithmetical consequences. The classical expansions are not in the appropriate form; we require the arithmetical developments in which the coefficients of the powers of q are given as explicit functions of the divisors of the exponents. These developments are unique. The chief elliptic theta series for paraphrase fall naturally into two sets according as they do or do not explicitly involve class numbers in their coefficients. This paper is concerned only with the latter kind and their more immediate paraphrases; but in all work like the present with the arithmetic of elliptic functions, these series of the first set appear to be fundamental, presenting themselves repeatedly in the most diverse investigations. Hence we shall give a fairly representative selection from them.

In writing down the few paraphrases of this paper we have aimed merely to show how such lists of properly prepared formulas may be used, much as a table of logarithms in other computations, for the almost immediate discovery of paraphrases broadly of the Liouville kind. We have purposely omitted all applications to specialized functions and their related theorems, the method of deriving special results being sufficiently evident from the papers of Liouville and Pepin, and from Bachmann's book. In connection with the series only brief notes on the calculations, in all cases simple, have been retained; but there are sufficient indications of the course followed for all the expansions to be quickly rechecked if desired.

2. The m, n, 2^a, d, δ, T_1, T_2, T_3 notation, explained in § 7 of Part I, is used throughout; and in the elliptic or theta series the summations refer to

[*] Read before the San Francisco Section of the American Mathematical Society, October, 1918. To save space, the extensive lists of theta and other expansions contained in the paper as read, have been omitted. Some of these, for the doubly periodic functions of the third kind, will appear elsewhere. Only the series necessary for illustrating the method of paraphrase in V have been retained.

[†] These Transactions, vol. 22 (1921), pp. 1–30.

all values > 0 of the type T_1, T_2 or T_3 indicated, the coefficients of the several powers of q being written after those powers in ( ), { } or [ ] as convenient. The ϑ_a notation is that of Jacobi (*Werke*, vol. 1, p. 501), with ϑ_0 in place of his ϑ. After glancing at the notation in §§ 3, 4 here and Part I § 7, the reader may pass at once to § 12, using §§ 3–11 only for reference. An important desideratum in regard to the series is pointed out in § 7, footnote.

3. *Numerical functions.* The constants occurring throughout this kind of work, other than class numbers, depend chiefly upon the functions now defined. Let $\zeta_r(n)$, $\zeta'_r(n)$, $\zeta''_r(n)$ denote respectively the sum of the rth powers of all, of the odd, of the even divisors of n; and let $\xi_r(n)$ denote the excess of the sum of the rth powers of all divisors $\equiv 1 \bmod 4$ of n over the sum of the rth powers of all those $\equiv -1 \bmod 4$; also let $\xi'_r(n)$ denote the excess of the sum of the rth powers of all divisors of n whose conjugates are $\equiv 1 \bmod 4$ over the sum of the rth powers of all those whose conjugates are $\equiv -1 \bmod 4$; and define $\xi''_r(n)$ by the identity $\xi'_r(n) + \xi''_r(n) = \xi_r(n)$. Write

$$\zeta_0(n),\ \zeta'_0(n),\ \zeta''_0(n),\ \xi_0(n),\ \xi'_0(n),\ \xi''_0(n) \equiv \zeta(n),\ \cdots,\ \xi''(n),$$

denoting the respective numbers of divisors pertaining to the six classes defined by the functions. For convenience we introduce six further functions

$$\alpha_r(n) = n^r\,\zeta'_{-r}(n); \qquad \lambda_r(n) = [1 + 2(-1)^n]\,\zeta'_r(n);$$

$$\mu_r(n) = \zeta_r(n) + \zeta'_r(n);$$

$$\beta_r(n) = 4\xi'_r(n) - \xi_r(n); \qquad \nu_r(n) = \zeta''_r(n) - (-1)^n\,\zeta'_r(n);$$

$$\rho_r(n) = \zeta'_r(n) - \zeta''_r(n);$$

and as before write $\alpha_0(n),\ \cdots,\ \rho_0(n) \equiv \alpha(n),\ \cdots,\ \rho(n)$. Of these, $\alpha_r(n)$ is seen to be the sum of the rth powers of all those divisors of n whose conjugates are odd. The equations for the rest express frequently occurring functions of the divisors which it is unnecessary at this point to define verbally. All twelve functions $\zeta_r,\ \cdots,\ \rho_r$ will be recognized as those which first present themselves in the principal theorems concerning representations of n as a sum of $2, 4, 6$ or 8 squares; and in the simplest applications of the paraphrases, such as those arising from $f(x|) = 1, x^2, x^4, \cdots, g(|x) = x, x^3, x^5, \cdots$ for all values of x, they reappear in many investigations, including those for $3, 5, 7, 9, 11$ or 13 squares.* In reductions of formulas the following immediate consequences of their definitions are most frequently useful.

* The treatment of these odd numbers of squares is given in the American Journal of Mathematics, vol. 42 (1920), pp. 168–188.

$$\begin{aligned} m &\equiv -1 \bmod 4, \qquad \xi(m) = 0; \\ m &\equiv \pm 1 \bmod 4, \qquad \xi_r'(m) = (-1)^{(m-1)/2}\,\xi_r(m); \end{aligned} \tag{1}$$

$\xi_r''(m) = 0$ or $2\xi_r(m)$ according as $m \equiv +1$ or $-1 \bmod 4$.

For $n = 2^\alpha m$, $m = d\delta$, $\alpha \geqq 0$; (2)–(7)

$$\begin{aligned} \zeta_r'(n) + \zeta_r''(n) = \zeta_r(n) &= [2^{(\alpha+1)r} - 1]\,\zeta_r(m)/(2^r - 1); \\ \zeta_r'(n) &= \zeta_r(m); \end{aligned} \tag{2}$$

$$(2^r - 1)\,\zeta_r''(n) = 2^r(2^{\alpha r} - 1)\,\zeta_r'(n); \tag{3}$$

$$2\sum 2^\alpha d = 2^{\alpha+1}\zeta_1'(n) = \zeta_1(n) + \zeta_1'(n) = \mu_1(n); \tag{4}$$

$$\zeta_1'(2n) = \zeta_1'(n); \qquad \zeta_1''(2n) = 2\zeta_1(n) = \tfrac{1}{2}\,\zeta_1''(2n) - \zeta_1'(2n); \tag{5}$$

$$\zeta_1(n) + \zeta_1''(n) = \zeta_1''(2n) - \zeta_1'(2n) = -\rho_1(2n); \tag{6}$$

$$\begin{aligned} \sum(-1)^{(d+\delta)/2} d &= (-1)^{(m+1)/2}\,\zeta_1(m); \\ \sum(-1)^{(d-1)/2} &= \xi(n) = \xi(m); \end{aligned} \tag{7}$$

and for $n = d\delta$,

$$\sum(-1)^{d+\delta} d = -\lambda_1(n); \qquad \sum(-1)^{d+\delta} = \zeta''(n) - \lambda(n). \tag{8}$$

To emphasize once more the notation, which will be followed in all subsequent lists, the Σ in (4), (7), (8) refers to all divisors d, δ of the indicated types, here T_2 for (4), (7), and T_3 for (8).

4. *Theta series and constants*; $\vartheta_\alpha(x) \equiv \vartheta_\alpha(x, q)$, $\vartheta_\alpha \equiv \vartheta_\alpha(0)$.

$$\vartheta_0(x) = 1 + 2\sum(-1)^n q^{n^2}\cos 2nx, \qquad \vartheta_0 = 1 + 2\sum(-1)^n q^{n^2}; \tag{9}$$

$$\begin{aligned} \vartheta_1(x) = 2\sum(-1)^{(m-1)/2} q^{m^2/4}\sin mx, \qquad \vartheta_1' &= \vartheta_0\,\vartheta_2\,\vartheta_3 \\ &= 2\sum(-1)^{(m-1)/2} mq^{m^2/4}; \end{aligned} \tag{10}$$

$$\vartheta_2(x) = 2\sum q^{m^2/4}\cos mx, \qquad \vartheta_2 = 2\sum q^{m^2/4}; \tag{11}$$

$$\vartheta_3(x) = 1 + 2\sum q^{n^2}\cos 2nx, \qquad \vartheta_3 = 1 + 2\sum q^{n^2}. \tag{12}$$

With but a few exceptions which can be derived from the others by means of the transformation of the second order, all of the series for $\vartheta_\alpha^a\,\vartheta_\beta^b$ in which $(\alpha, \beta) = (2, 3)$, $(0, 3)$, $(0, 2)$ and $(a, b) = (1, 1)$, $(2, 2)$, $(2, 4)$, $(4, 2)$, $(3, 3)$, $(4, 4)$, $(0, 2)$, $(0, 4)$, $(0, 6)$, $(0, 8)$, can be simply found from the series for the k, k', K constants and their powers as given by Jacobi in §§ 40–42 of the *Fundamenta Nova.* For the rapid and systematic use of the method of paraphrase the series for all of these constants will be found indispensable. The coefficients of all are expressible directly in terms of the numerical functions defined in § 3. Here however we need give only the following selection:

$$(13)\quad \begin{gathered} \vartheta_2\,\vartheta_3 = 2\sum q^{m/4}\,\xi(m); \qquad \vartheta_0\,\vartheta_3 = 1 + 4\sum q^{2n}(-1)^n\,\xi(n); \\ \vartheta_0\,\vartheta_2 = 2\sum q^{m/4}(-1)^{(m-1)/4}\,\xi(m); \end{gathered}$$

$$(14)\quad \begin{gathered} \vartheta_0^2 = 1 + 4\sum q^n(-1)^n\,\xi(n); \qquad \vartheta_2^2 = 4\sum q^{m/2}\,\xi(m); \\ \vartheta_3^2 = 1 + 4\sum q^n\,\xi(n); \end{gathered}$$

and the following essential constants in reduced form from Jacobi, *l. c.*, § 41:

$$(15)\quad \begin{gathered} A = 4\sum q^n\,\mu_1(n); \qquad B = -4\sum q^n(-1)^n\,\mu_1(n); \\ C = 1 + 8\sum q^{2n}\,\rho_1(n); \end{gathered}$$

$$(16)\quad \begin{gathered} B + C = 1 - 8\sum q^n(-1)^n\,\nu_1(n); \qquad A - B = 16\sum q^{2n}\,\mu_1(n); \\ C - A = 1 - 8\sum q^n\,\nu_1(n), \end{gathered}$$

of which C is $4KE^1/\pi^2$, and all are required in the derivations of the series in § 6.

5. *Eighteen doubly periodic theta quotients.** We shall give explicitly only those of the eighteen which can not be derived from others by changes of q into $-q$ and x into $x + \pi/2$, or by the transformation of the second order.

$$(17)\quad \phi_1(x) = \vartheta_2\,\vartheta_3\,\vartheta_1(x)/\vartheta_0(x) = 4\sum q^{m/2}\left(\sum \sin dx\right),\ [T_1];$$

$$(18)\quad \begin{aligned} \phi_2(x) &= \vartheta_2\,\vartheta_3\,\vartheta_0(x)/\vartheta_1(x) \\ &= \csc x + 4\sum q^n\left[\sum(-1)^{(d-1)/2}\cos dx\right],\ [T_2]; \end{aligned}$$

$$(19)\quad \begin{aligned} \phi_9(x) &= \vartheta_0\,\vartheta_3\,\vartheta_3(x)/\vartheta_0(x) \\ &= 1 + 4\sum q^n\left[\sum(-1)^{(\delta-1)/2}\cos 2^{a+1}dx\right],\ [T_2]; \end{aligned}$$

$$(20)\quad \begin{aligned} \phi_{11}(x) &= \vartheta_0\,\vartheta_3\,\vartheta_2(x)/\vartheta_1(x) \\ &= \cot x + 4\sum q^{2n}\left[\sum(-1)^{\delta}\sin 2dx\right],\ [T_3]; \end{aligned}$$

in which, *as always henceforth*, the $[T_a]$ indicates the type of division for the m or n in the exponent. Note particularly that when the exponent is cn, c being a numerical constant, the type refers to the divisors of n, and not of cn, and so in all similar cases.

It will be sufficient here to indicate how the remaining fourteen functions in Hermite's list (which includes all of the doubly periodic functions considered by Jacobi, and six others), may be derived from these. In (21) the first

* Calculated (with corrections) from the equivalent set given by Hermite, *Sur les Théorèmes de M. Kronecker, etc.*, *Oeuvres*, vol. 1, p. 243, or Journal de Mathématiques pures et appliquées, (2), vol. 9 (1864), p. 145. The suffixes correspond to the order of the functions in Hermite's list. For the connection of these with the doubly periodic functions of the second kind (§ 10), cf. Messenger of Mathematics, vol. 49 (1919–20), p. 81.

member in each triad of functions is transformed into the second by the substitution $x \sim x + \pi/2$, and into the third by $q \sim -q$; $i \equiv \sqrt{-1}$. Thus, carrying out the indicated transformations we find from (17) and the first triad the developments of

$$\phi_3(x) = \vartheta_2\, \vartheta_3\, \vartheta_2(x)/\vartheta_3(x)$$

upon changing x into $x + \pi/2$ in (17), and

$$\phi_5(x) = \vartheta_0\, \vartheta_2\, \vartheta_1(x)/\vartheta_3(x)$$

on replacing q by $-q$ in (17) and reducing. Similarly for all fourteen.

$$
\begin{array}{lllll}
 & (\phi_1, \phi_3, i\phi_5), & (\phi_3, -\phi_1, i\phi_7), & (\phi_7, -\phi_5, i\phi_3); & \\
 & (\phi_2, \phi_4, \phi_6), & (\phi_4, -\phi_2, \phi_8), & (\phi_6, \phi_8, \phi_2), & (\phi_8, -\phi_6, \phi_4); \\
(21) & (\phi_9, \phi_{10}, \phi_{10}), & (\phi_{10}, \phi_9, \phi_9), & (\phi_{11}, -\phi_{12}, \phi_{11}), & (\phi_{12}, -\phi_{11}, \phi_{12}); \\
 & (\phi_{13}, -\phi_{14}, \phi_{15}), & (\phi_{14}, -\phi_{13}, \phi_{16}), & (\phi_{15}, -\phi_{16}, \phi_{13}), & (\phi_{16}, -\phi_{15}, \phi_{14}); \\
 & (\phi_{17}, -\phi_{17}, -\phi_{17}), & (\phi_{18}, -\phi_{18}, \phi_{18}); & &
\end{array}
$$

with which we need for the last two,

$$(22) \qquad \phi_{17}(x, q) = 2\phi_1(2x, q^2), \qquad \phi_{18}(x, q) = 2\phi_2(2x, q^2),$$

which, upon reducing the functions on the right by the transformation of the second order to functions of x, q, yield Hermite's forms. It may be noted that the set yields no further quotients of the same kind under any of these transformations.

The derivation of the expansions (17)–(20) is sufficiently evident from Hermite's detailed similar calculation in the supplement to Lacroix' *Calculus*, reproduced in *Oeuvres*, vol. 1, pp. 219–220. Hermite's remarks regarding these fundamental developments are so apposite in the present connection that we quote them. "Il est impossible de ne pas être frappé du caractère arithmétique de ces expressions (Σ sin dx, etc.); elles offrent un exemple des fonctions numériques qui ont été le sujet des belles recherches de M. Liouville, et la manière simple dont elles sont amenées par la théorie des fonctions elliptiques peut aisément faire présumer le rôle de cette théorie dans l'étude des propriétés des nombres." It was from precisely this observation that the (probably) true origin of Liouville's general formulas first became evident: in fact his first is equivalent to the identity $[\phi_1(x)]^2 = \phi_1(x) \times \phi_1(x)$. Others of his formulas (of which we shall omit discussion) follow from equally simple identities, such as

$$\phi_1(x)\phi_2(x) = \phi_3(x)\phi_4(x) = \vartheta_2^2\, \vartheta_3^2; \qquad \phi_{15}(x)\phi_{16}(x) = \vartheta_3^4;$$

$\phi_1' = \phi_7\, \phi_9$, where ϕ_1' denotes the x-derivative* of $\phi_1(x)$.

* For the complete set of these in reduced form, cf. Messenger of Mathematics, vol. 47 (1917–18), p. 55, where are also given the calculations for § 6.

6. *Squares of the eighteen* $\phi_j(x)$. All may be inferred from the following by means of (21), (22).

$$(23)\quad \phi_1^2(x) = 4\sum q^n \mu_1(n) - 8\sum q^n (\sum 2^a d \cos 2^{a+1} dx), \; [T_2];$$

$$(24)\quad \phi_2^2(x) = 4\sum q^n \mu_1(n) + \csc^2 x - 8\sum q^{2n} (\sum d \cos 2dx), \; [T_3];$$

$$(25)\quad \phi_9^2(x) = 1 + 8\sum q^{2n} \rho_1(n) + 8\sum q^n (\sum 2^a d \cos 2^{a+1} dx), \; [T_2];$$

$$(26)\quad \phi_{11}^2(x) = -8\sum q^{2n} \rho_1(n) + \cot^2 x - 8\sum q^{2n} (\sum d \cos 2dx), \; [T_3].$$

For the verification of these, (15), (16) will be found necessary.

7. In explanation of the unusual forms of these expansions which, if considered only for their use in the customary applications of elliptic functions are of slight significance, it will be well to state here what are the desirable characteristics, from the standpoint of paraphrases, for such series and identities between them to possess. Consider for example

$$2\vartheta_0^4(x)\phi_1(x)\phi_7(x)\phi_9(x) = \vartheta_1'^3\,\vartheta_1(2x),$$

which is easily seen to be true. On the left we have the product of seven known series, viz., $\vartheta_0(x)$ counted four times, by $\phi_1(x)$, $\phi_7(x)$, $\phi_9(x)$; while on the right we have the product of only four, or, if the transformation of the third order be used to give the series for $\vartheta_1'^3$, only two. Or, the left may clearly be read as the product of six in three ways, each of $\phi_1\phi_7$, $\phi_7\phi_9$, $\phi_9\phi_1$ being known from the derivative $\phi_1'(x)$. Hence, reading the identity in the first way, we shall get a paraphrase connecting *L*-functions of parity $(0|1)$ integrated over two separations of degrees 7 and 4 respectively, or 7 and 2; in the other ways the degrees are 6, 4 (three times) and 5, 4. Therefore since in these the separations of degree 4 or 2 are the same, we have a syzygy between integrations over separations of degrees 4 (or 2), 5, 6, 7. Thus it follows that these arithmetical facts connected with the separations of degrees 7, 6, 5 (which relate respectively to representations in quadratic forms containing 7, 6, 5 indeterminates), may be reduced to others concerning separations of degrees 4 or 2 (which relate to representations in quaternary or binary quadratic forms), and the theorems corresponding to the specializations of the *L*-functions, similarly reduced. This course has obviously been followed, for simpler identities, by Liouville in his 17th and 18th memoirs; also elsewhere. Our object then is to find, where possible, simple expansions for fairly complicated functions, in order to reduce complex arithmetical relations to others which are simpler. The greater the complexity of the function which is reduced to a simpler product of known series, and itself expanded in a power series in q, the greater will be the variety and interest of the paraphrases. The most desirable case is the reduction of a product of many theta factors to a single series whose coefficients are given as explicit functions of the real divisors of the exponents of q.*

8. All possible cases of the paraphrase of identities involving tangents, cotangents, secants or cosecants are covered by the sixteen formulas following. They may be verified by inspection on multiplying throughout by $\sin x$ in the cases of the cotangents and cosecants, by $\cos x$ in the others. The suffix in $u_0(x), \cdots, w_7(x)$ is even or odd according as the integer in ( ) is even or odd; and where necessary to indicate both variables, the functions will be written $u_0(2n, x), \cdots, w_7(2n, x)$.

$$\begin{aligned}
u_0(2n) &\equiv \sin 2nx \csc x = 2\sum_{r=1}^{n} \cos(2r-1)x,\\
u_1(m) &\equiv \sin mx \csc x = 1 + 2\sum_{r=1}^{(m-1)/2} \cos 2rx,\\
u_2(2n) &\equiv \sin 2nx \sec x = 2(-1)^n \sum_{r=1}^{n} (-1)^r \sin(2r-1)x,\\
u_3(m) &\equiv \sin mx \sec x = (-1)^{(m-1)/2}\left[\tan x + 2\sum_{r=1}^{(m-1)/2} (-1)^r \sin 2rx\right].
\end{aligned}\tag{27}$$

$$\begin{aligned}
v_0(2n) &\equiv \cos 2nx \csc x = \csc x - 2\sum_{r=1}^{n} \sin(2r-1)x,\\
v_1(m) &\equiv \cos mx \csc x = \cot x - 2\sum_{r=1}^{(m-1)/2} \sin 2rx,\\
v_2(2n) &\equiv \cos 2nx \sec x\\
&\qquad = (-1)^n\left[\sec x + 2\sum_{r=1}^{n} (-1)^r \cos(2r-1)x\right],\\
v_3(m) &\equiv \cos mx \sec x = (-1)^{(m-1)/2}\left[1 + 2\sum_{r=1}^{(m-1)/2} (-1)^r \cos 2rx\right].
\end{aligned}\tag{28}$$

$$\begin{aligned}
w_0(2n) &\equiv \sin 2nx \cot x = 1 + \cos 2nx + 2\sum_{r=1}^{n-1} \cos 2rx,\\
w_1(m) &\equiv \sin mx \cot x = \cos mx + 2\sum_{r=1}^{(m-1)/2} \cos(2r-1)x,\\
w_2(2n) &\equiv \sin 2nx \tan x\\
&\qquad = (-1)^{n-1}\left[1 + (-1)^n \cos 2nx + 2\sum_{r=1}^{n-1} (-1)^r \cos 2rx\right],\\
w_3(m) &\equiv \sin mx \tan x\\
&\qquad = (-1)^{(m-1)/2}\left[\sec x + (-1)^{(m+1)/2} \cos mx\right.\\
&\qquad\qquad \left. + 2\sum_{r=1}^{(m-1)/2} (-1)^r \cos(2r-1)x\right].
\end{aligned}\tag{29}$$

* In particular much light would be thrown on the arithmetic of quadratic forms in n indeterminates by the corresponding developments of

$$\vartheta_\alpha(x_1 + x_2 + \cdots + x_n)/\vartheta_\beta(x_1)\vartheta_\gamma(x_2)\cdots\vartheta_\delta(x_n),$$

$$
\begin{aligned}
&w_4(2n) \equiv \cos 2nx \cot x = \cot x - \sin 2nx - 2\sum_{r=1}^{n-1} \sin 2rx,\\
&w_5(m) \equiv \cos mx \cot x = \csc x - \sin mx - 2\sum_{r=1}^{(m-1)/2} \sin(2r-1)x,\\
&w_6(2n) \equiv \cos 2nx \tan x\\
(30)\qquad &\qquad = (-1)^n\left[\tan x + (-1)^n \sin 2nx + 2\sum_{r=1}^{n-1}(-1)^r \sin 2rx\right],\\
&w_7(m) \equiv \cos mx \tan x\\
&\qquad = \sin mx - 2(-1)^{(m-1)/2}\sum_{r=1}^{(m-1)/2}(-1)^r \sin(2r-1)x.
\end{aligned}
$$

(28)-(30) are connected by many simple and interesting group relations which, as they lie off the main course of this paper, we omit. We may notice, however, a theorem of which the proof presents no difficulty, and which often either gives additional paraphrases or affords a check on the reductions of theta series. Let $h(x)$ denote any one of $\sec x$, $\csc x$, $\tan x$, $\cot x$, and A, B quantities independent of x. Suppose that over some separation the following is an identity in x,

$$Ah(x) + B + \sum a_i t(n_i x) = 0,$$

in which t represents either sin or cos. Then this implies

$$A = 0, \qquad B + \sum a_i t(n_i x) = 0.$$

9. We frequently meet expressions involving secants, etc., of several variables to be paraphrased. An example will make clear the procedure in all such cases. Writing $z = x + y$, we have

$$
\begin{aligned}
\sin(mx + 2ny)\csc(x+y) &= \sin\{(m-2n)x + 2nz\}\csc z\\
&= v_0(2n, z)\sin(m-2n)x + u_0(2n, z)\cos(m-2n)x,
\end{aligned}
$$

and each term is in a form suitable for paraphrase.

10. *Doubly periodic functions of the second kind.** From the standpoint of paraphrase, these functions are of the highest importance. They not only

where $\alpha, \beta, \gamma, \cdots, \delta$, are any of the numbers $0, 1, 2, 3$. The case $n = 2$ is considered below, § 11. Note that, as pointed out by Glaisher (Messenger of Mathematics, vol. 14 (1884–85), p. 162), these, the arithmetical developments, are unique, while the analytical representations are not.

* The nomenclature is that of Hermite. The doubly periodic functions of the third kind also are of great use in this subject. They are particularly valuable in the derivation of new and generalized class number relations. As the forms of these functions due to Hermite, Biehler, Appel and others can be changed to their arithmetical forms only after many reductions, we shall give the appropriate developments elsewhere.

include the doubly periodic functions as limiting cases[*] in one direction and give rise to a great variety of doubly periodic functions of the third kind in another, but they also afford us the first and most essential link connecting paraphrases in which the order of the functions is unity with those in which the order exceeds unity. It is of interest to note that equivalents of the series given presently can easily be derived from Jacobi's investigations on the rotation of a rigid body and were, therefore, most probably familiar to Liouville.

Consider the function

$$\phi_{\alpha\beta\gamma}(x_1, x_2) = \vartheta_1' \, \vartheta_\alpha(x_1 + x_2)/\vartheta_\beta(x_1)\,\vartheta_\gamma(x_2), \equiv \phi_{\alpha\beta\gamma};$$

and denote by $S_i \phi_{\alpha\beta\gamma}$, $P_j \phi_{\alpha\beta\gamma}$ $(i, j = 1, 2)$ the results of replacing x_i, x_j by $x_i + \pi/2$, $x_j + \pi\tau/2$ respectively, where τ has the usual meaning in terms of the half-periods; and let $\cdots P_i P_j S_i P_j$ denote the application of the substitutions $P_j, S_i, P_j, P_i, \cdots$ in the order last written. Construct the substitutions (which do not form a group) σ_i $(i = 1, 2, \cdots, 16)$:

$$\begin{array}{llll}
\sigma_1 = 1, & \sigma_2 = S_2, & \sigma_3 = S_1 S_2, & \sigma_{11} = S_1; \\
\sigma_4 = P_2, & \sigma_5 = S_2 P_2, & \sigma_6 = S_1 P_2, & \sigma_7 = S_2 S_1 P_2; \\
\sigma_{12} = P_1, & \sigma_{13} = S_1 P_1, & \sigma_{14} = S_2 P_1, & \sigma_{15} = S_2 S_1 P_1; \\
\sigma_8 = P_1 P_2, & \sigma_9 = S_2 P_1 P_2, & \sigma_{10} = S_2 S_1 P_1 P_2, & \sigma_{16} = S_1 P_1 P_2;
\end{array}$$

and apply them to ϕ_{100}, giving:[†]

$$\begin{array}{llll}
\phi_1 = \phi_{100}, & \phi_2 = \phi_{203}, & \phi_3 = \phi_{133}, & \phi_{11} = \phi_{230}; \\
\phi_4 = \phi_{001}, & \phi_5 = \phi_{302}, & \phi_6 = \phi_{331}, & \phi_7 = \phi_{032}; \\
\phi_{12} = \phi_{010}, & \phi_{13} = \phi_{320}, & \phi_{14} = \phi_{313}, & \phi_{15} = \phi_{023}; \\
\phi_8 = \phi_{111}, & \phi_9 = \phi_{212}, & \phi_{10} = \phi_{122}, & \phi_{16} = \phi_{221}.
\end{array}$$

It may be verified that, disregarding signs, these sixteen functions are all those that can be generated from any one of them, ϕ_k, by successive applications of S_1, S_2, P_1, P_2. Hence from the series for ϕ_k may be written down those for the remaining fifteen, and no others, by this process of transformation. The paraphrase interpretation of S_1, S_2 has been considered in Part I, § 30; that of P_1, P_2 is by no means so obvious, and need not detain us here.

[*] Messenger of Mathematics, vol. 49 (1919–20), p. 81.

[†] The factor ± 1 being immaterial for our purpose, it is ignored. The sign is − for $j = 3$, 10, and for the rest $+$. In all there are 64 possible functions $\phi_{\alpha\beta\gamma}$; the remaining 48 need not concern us here. There is an obvious advantage in deriving the functions from one of them, ϕ_{100} (any other of the 16 might have been selected), instead of from two distinct fundamental series as done by Hermite.

11. *Series for the functions of the second kind.** Applying the substitutions σ to ϕ_{100}, writing $\phi_{\alpha\beta\gamma} \equiv \phi_{\alpha\beta\gamma}(x, y)$, x', $y' = x + \pi/2$, $y + \pi/2$, and developing in powers of q, we find after all reductions the following four in which the type is T_1,

$$\phi_1(x, y) \equiv \phi_{100} = 4\sum q^{m/2}[\sum \sin(dx + \delta y)], \tag{31}$$

$$\phi_2(x, y) \equiv \phi_{203} = \phi_1(x, y') = 4\sum q^{m/2}[\sum(-1)^{(\delta-1)/2}\cos(dx + \delta y)], \tag{32}$$

$$\begin{aligned}\phi_3(x, y) \equiv \phi_{133} &= -\phi_1(x', y') \\ &= 4\sum q^{m/2}[(-1)^{(m-1)/2}\sum \sin(dx + \delta y)],\end{aligned} \tag{33}$$

$$\phi_{11}(x, y) \equiv \phi_{230}(x, y) = \phi_2(y, x); \tag{34}$$

eight in which the type is T_2,

$$\phi_4(x, y) \equiv \phi_{001} = \csc y + 4\sum q^n[\sum \sin(2^{\alpha+1} dx + \delta y)], \tag{35}$$

$$\begin{aligned}\phi_5(x, y) \equiv \phi_{302} &= \phi_4(x, y') \\ &= \sec y + 4\sum q^n[\sum(-1)^{(\delta-1)/2}\cos(2^{\alpha+1} dx + \delta y)],\end{aligned} \tag{36}$$

$$\begin{aligned}\phi_6(x, y) \equiv \phi_{331} &= \phi_4(x', y) \\ &= \csc y + 4\sum q^n[(-1)^n \sum \sin(2^{\alpha+1} dx + \delta y)],\end{aligned} \tag{37}$$

$$\begin{aligned}\phi_7(x, y) \equiv \phi_{032} &= \phi_4(x', y') \\ &= \sec y + 4\sum q^n[(-1)^n \sum(-1)^{(\delta-1)/2}\cos(2^{\alpha+1} dx + \delta y)],\end{aligned} \tag{38}$$

$$\phi_{12}(x, y) \equiv \phi_{010} = \phi_4(y, x), \tag{39}$$

$$\phi_{13}(x, y) \equiv \phi_{320} = \phi_5(y, x), \tag{40}$$

$$\phi_{14}(x, y) \equiv \phi_{313} = \phi_6(y, x), \tag{41}$$

$$\phi_{15}(x, y) \equiv \phi_{023} = \phi_7(y, x); \tag{42}$$

and four in which the type is T_3,

* This list being of such importance it was calculated and checked in several ways, to eliminate printers' errors prevalent in other forms in the literature. It was calculated: (i) by Hermite's method, *Sur quelques applications des fonctions elliptiques* (Comptes Rendus, vol. 85 (1877), $\cdots$ 94 (1882), *Oeuvres*, vol. 3, p. 267; (ii) by applying the σ_j to Halphen's form of $\phi_8(x, y)$ (*Traité*, vol. 1, p. 418 (15)), and comparing with the same for $\phi_1(x, y)$; (iii) by carrying out in detail the calculations in Hermite's paper *Sur une application de la théorie des fonctions doublement périodiques de seconde espèce*, (Annales de l'École Normale Supérieure, (3), vol. 2 (1885), p. 303, reprinted with corrections in *Oeuvres*, vol. 4, p. 190). Finally it was compared with Hermite's final list, *Oeuvres*, vol. 4, pp. 199–200, which still contains an error (in the expansion of $\phi_{10}(x, y)$, as may be verified upon putting $x, y = 0, \pi/2$ and comparing with the series for the elliptic functions). Equivalent forms for certain members of this list quoted by other writers with an indefinite reference to Kronecker, are unreliable, and should be recalculated before use.

$$\phi_8(x, y) \equiv \phi_{111} = \cot x + \cot y + 4\sum q^{2n}\left[\sum \sin 2(dx + \delta y)\right], \tag{43}$$

$$\begin{aligned} \phi_9(x, y) \equiv \phi_{212} = \phi_8(x, y') = \cot x - \tan y \\ + 4\sum q^{2n}\left[\sum (-1)^{\delta} \sin 2(dx + \delta y)\right], \end{aligned} \tag{44}$$

$$\begin{aligned} \phi_{10}(x, y) \equiv \phi_{122} = -\phi_8(x', y') = \tan x + \tan y \\ - 4\sum q^{2n}\left[\sum (-1)^{d+\delta} \sin 2(dx + \delta y)\right], \end{aligned} \tag{45}$$

$$\phi_{16}(x, y) \equiv \phi_{221} = \phi_9(y, x). \tag{46}$$

V. The method of paraphrase, simple illustrations *

12. As already remarked, lists of formulas in the d, δ form, such as the foregoing, are analogous in paraphrasing to tables of logarithms in common arithmetic. It will be evident that by combining the series to form identities there is implicit in the lists given an infinity of paraphrases such as those exemplified in Part I; and as a systematic derivation of all the most obvious paraphrases is out of the question in a paper of this length, we shall limit the illustrations to a few only of the paraphrases lying on the surface, choosing the examples partly for their own interest, and partly to show one or two of the simpler methods for using such lists of developments. For this purpose we may select L-functions of degrees $1, 2$, and confine our attention principally to linear separations, the other cases being treated with equal facility.

13. By the theory developed in Part I, trigonometric products are always to be written in their equivalent sum forms before proceeding to paraphrase. We shall accordingly write trigonometric identities derived from elliptic in the latter form at once, omitting the intermediate product forms, unless the separation into sums is not obvious.

As a first example, consider the eighteen identities of the form

$$\phi_j^2(x) = \phi_j(x) \times \phi_j(x).$$

From (17), (23),

$$\phi_1(x, q^2) \times \phi_1(x, q^2) = \phi_1^2(x, q^2)$$

is equivalent to

$$\begin{aligned} 16\sum q^m\left(\sum \sin dx\right) \times \sum q^m\left(\sum \sin dx\right) \\ = 4\sum q^{2n}\mu_1(n) - 8\sum q^{2n}\left[2^a \sum d \cos 2^{a+1} dx\right], \end{aligned}$$

where, *as in all such cases*, the types of division being defined in the lists from which the series are transcribed, need not be written; here they are T_1 on the

* At this point it will be advantageous to glance through the Introduction to Part I, attending particularly to the bar notation for L-functions and the notation for separations there explained, also to the illustrative examples of paraphrases, as we shall not again refer to any of the notation. See also the references at the end of Part I. We shall cite Liouville's theorems by giving the numbers of the memoir, page and formula, thus: Liouville (1), 144, (A).

left, T_2 on the right. On referring to § 3 (4) for the d-form of $\mu_1(n)$, and equating coefficients of q^{2n}, we get

$$2n = m' + m''; \qquad n = 2^\alpha m; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$
$$\sum [\cos(d' - d'')x - \cos(d' + d'')x] = 2^\alpha \sum d[1 - \cos 2^{\alpha+1} dx];$$

and this, on writing $f(x|) \equiv f(x)$, paraphrases into

$$\text{(I)} \quad \begin{gathered} 2n = m' + m''; \qquad n = 2^\alpha m; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'': \\ \sum [f(d' - d'') - f(d' + d'')] = 2^\alpha \sum d[f(0) - f(2^{\alpha+1} d)]; \end{gathered}$$

which is Liouville's (2), 194, (a), and for $\alpha = 0$ his first formula, (1), 144, (A). It is of course not necessary in any such case as this to replace the numerical functions $\mu_1(n)$, etc., by their d-forms; but doing so increases the symmetry of the paraphrases. For simplicity in writing we shall *henceforth put*

$$f(x|) \equiv f(x).$$

In (I) each argument is even. Hence $f(x)$ may be replaced by $f(x/2)$, giving

$$\sum \left[f\left(\frac{d' - d''}{2}\right) - f\left(\frac{d' + d''}{2}\right)\right] = 2^\alpha \sum d[f(0) - f(2^\alpha d)].$$

In this case no material simplification is thus effected. But when the contrary is the case we shall make the change without notice. A similar transformation of a paraphrase involving odd arguments is not permissible, since, by the definition of an L-function, $f(m/2)$ does not necessarily exist.

14. The use of (27)–(30) will be sufficiently clear from the following, which we give in some detail as it illustrates transformations of several types which occur frequently. Any identity involving csc, tan or cot might be chosen; we take $\phi_2(x) \times \phi_2(x) = \phi_2^2(x)$, and use (18), (24), getting

$$[\csc x + 4\sum q^n (\sum \sin dx)]^2 = 4\sum q^n \mu_1(n) + \csc^2 x - 8\sum q^{2n} (\sum d \cos 2dx).$$

Equating coefficients of q^n, where $n = 2^\alpha m$, $\alpha > 0$, we have for the separations $n/2 = d_1 \delta_1$, and

$$n = n' + n''; \qquad n, n', n'' = 2^\alpha m, 2^{\alpha'} m', 2^{\alpha''} m'';$$
$$m, m', m'' = d\delta, d'\delta', d''\delta'',$$

the following

$$\sum \sin dx \csc x + \sum [\cos(d' - d'')x - \cos(d' + d'')x]$$
$$= \tfrac{1}{2}\mu_1(n) - \sum d_1 \cos 2d_1 x.$$

Since d is odd, (27) gives for $\Sigma \sin dx \csc x$ the value $\Sigma u_1(d, x)$, or

$$\sum\left[1 + 2\sum_{r=1}^{(d-1)/2} \cos 2rx\right],$$

$$\equiv \zeta'(n) + 2\sum[\cos 2x + \cos 4x + \cos 6x + \cdots + \cos(d-1)x],$$

where we have taken ζ' from § 3. We may eliminate the separation $n/2 = d_1\,\delta_1$, by the following obvious identity

$$\sum d_1 \cos 2d_1 x \equiv \sum[d \cos 2dx + 2d \cos 2^2 dx + \cdots + 2^{\alpha-1} d \cos 2^\alpha dx],$$

which results upon segregating the odd divisors d and (when $\alpha > 1$) the even divisors $2^\beta d\,(0 < \beta \leqq \alpha - 1)$ among all the divisors d_1 of $2^{\alpha-1} m$. Hence, substituting the sums thus found for the respective terms of the original sin, cos, csc identity, and paraphrasing, we have

$$n = 2^\alpha m = 2^{\alpha'} m' + 2^{\alpha''} m''; \qquad \alpha > 0; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\begin{aligned} &\sum[f(d'-d'') - f(d'+d'')] = [\tfrac{1}{2}\mu_1(n) - \zeta'(n)]f(0) \\ \text{(II)}\qquad &\qquad - 2\sum[f(2) + f(4) + f(6) + \cdots + f(d-1)] \\ &\qquad - \sum d[f(2d) + 2f(2^2 d) + \cdots + 2^{\alpha-1} f(2^\alpha d)]. \end{aligned}$$

The subcases of all paraphrases such as (II) in which $\alpha = 0$ (although this does not directly come under (II), α being > 0 therein, the paraphrases are closely related, both being consequences of the $\phi_2(x)$ identity), are of great importance in connection with the representations of primes p in the form $ax^2 + br^c y^2$, where a, b are constants and r is prime, in that the specialized forms of these subcases for $f(x) = 1, x^2, \cdots$ give rise to identities of the forms considered by Bouniakowsky and Liouville as the point of departure for determining the number of such representations. In the present instance, the sum $\Sigma q^{2n}(\Sigma d_1 \cos 2d_1 x)$ can contribute nothing to the coefficient of q^m; hence the second Σ on the right of (II) is absent. Again, for $\alpha = 0$, we have $m = 2^{\alpha'} m' + 2^{\alpha''} m''$. Hence one and only one of $\alpha', \alpha'' = 0$; and therefore the value of $\Sigma[f(d'-d'') - f(d'+d'')]$ over the separation $n = 2^{\alpha'} m' + 2^{\alpha''} m''$ is the sum of its values over the (identical) separations

$$m = m' + 2^{\alpha''} m'', \qquad m = 2^{\alpha'} m' + m''.$$

Finally then we have, on referring to § 3 (4) for the numerical functions,

$$m = 2^{\alpha'} m' + m''; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\begin{aligned} \text{(III)}\quad &2\sum[f(d'-d'') - f(d'+d'')] = [\zeta_1(m) - \zeta(m)]f(0) \\ &\qquad - 2\sum[f(2) + f(4) + \cdots + f(d-1)]. \end{aligned}$$

This is Liouville's (3), 201, (D). We have not yet exhausted the obvious consequences of $\phi_2^2 = \phi_2 \times \phi_2$; any identity involving csc x gives at once two paraphrases, one similarly to (II), and the other by paraphrasing the result of multiplying the identity throughout by $\sin x$ to eliminate $\csc x$. In the present case we get, on equating coefficients of q^n, $n = 2^\alpha m$, $\alpha > 0$, as in proving (II):

$$\begin{aligned}\sum [\sin (1 + d' - d'')x + \sin (1 - d' + d'')x \\ + \sin (-1 + d' + d'')x - \sin (1 + d' + d'')x] \\ = \mu_1(n) \sin x - 2\sum \sin dx \\ - 2\sum \left[\sum_{r=1}^{\alpha} 2^{r-1} d\{\sin (2^r d + 1)x - \sin (2^r d - 1)x\} \right],\end{aligned}$$

which, on writing $g(x) \equiv f(\,|x)$, gives the paraphrase

$$n = 2^\alpha m = 2^{\alpha'} m' + 2^{\alpha''} m''; \qquad \alpha > 0; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(IV)} \quad \begin{aligned}\sum [g(1 + d' - d'') + g(1 - d' + d'') + g(-1 + d' + d'') \\ - g(1 + d' + d'')] = \mu_1(n)g(1) - 2\sum g(d) \\ - 2\sum \left[\sum_{r=1}^{\alpha} 2^{r-1} d\{g(2^r d + 1) - g(2^r d - 1)\} \right].\end{aligned}$$

The diversity in form of (II), (IV), is the more striking in that (IV) is merely the paraphrase of the same elliptic identity as that which gives rise to (II) when it is multiplied throughout by $\sin x$. For identities involving $\csc x$ or $\cot x$ we thus get the paraphrases corresponding to (IV) on multiplying first by $\sin x$; for those containing $\sec x$ or $\tan x$ we first multiply throughout by $\cos x$. Similarly to (IV), corresponding to (III), we get,

$$m = 2^{\alpha'} m' + m''; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(V)} \quad \begin{aligned}\sum [g(1 + d' - d'') + g(1 - d' + d'') + g(-1 + d' + d'') \\ - g(1 + d' + d'')] = \zeta_1(m)g(1) - \Sigma g(d),\end{aligned}$$

which is Liouville (3), 206, (E). He remarks that (III), (V) are ultimately the same thing, which is obvious from their origin. His (F), ibid., p. 208, is a paraphrase (among others) of Jacobi's $Z^2(u)$, *Fundamenta Nova*, § 47, equation (1); also his (4), 242, (G) is from the same source, or it follows from one of the formulas for functions $f(x, y\,|)$ given below in (XVII).

15. Although their arithmetical consequences are often widely different, at least in appearance, we shall regard paraphrases that may be derived from one another by means of the elementary transformations of Part I (III) as equivalent. We shall now show how the simplest properties of the elliptic or theta functions such as those in (21) are of direct use in finding all the

distinct and equivalent paraphrases implicit in a complete set of identities of a given kind, here $\phi_j(x) \times \phi_j(x) = \phi_j^2(x)$.

Obviously $\phi_{17}(x) \times \phi_{17}(x) = \phi_{17}^2(x)$ gives nothing distinct from (I), as may be seen on glancing at (22). In the same way we dispose of $j = 13, 14, 18$ when the paraphrase for $j = 11$ is known; and (21), (22) show that for the following pairs of values of j the $\phi_j^2 = \phi_j \times \phi_j$ paraphrases will be identical,

$$(1, 5),\ (2, 6),\ (3, 7),\ (4, 8),\ (9, 10),\ (13, 15),\ (14, 16),$$

since in each case the resulting elliptic identities may be transformed into one another by changing the sign of q. And from the same source it is seen that the paraphrases corresponding to the next pairs will be equivalent in the sense that either in a given pair may be transformed into the other by one of the elementary transformations considered in Part I (III),

$$(1, 3),\ (2, 4),\ (5, 7),\ (6, 8),\ (9, 10),\ (11, 12),\ (13, 14),\ (15, 16).$$

Hence we shall find all the required paraphrases by taking $j = 1, 2, 3, 4, 9, 11, 12$. The cases $j = 1, 2$ give the paraphrases (I)–(V); omitting the alternative forms that correspond to (IV), (V) we get the following in the same way for $j = 3, 4, 9, 11, 12$.

$$2n = m' + m''; \qquad n = 2^\alpha m; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(VI)} \quad \begin{aligned} &\sum [(-1)^{(d'+d'')/2} \{f(d' - d'') + f(d' + d'')\}] \\ &\qquad = 2^\alpha \sum d[(-1)^n f(2^{\alpha+1} d) - f(0)]; \\ &\sum [(-1)^{(\delta'+\delta'')/2} \{f(d' - d'') + f(d' + d'')\}] \\ &\qquad = -2^\alpha \sum d[f(0) + f(2^{\alpha+1} d)]. \end{aligned}$$

These are identical; the first comes from $j = 3$, the second from $j = 7$. For $j = 4$ we use r_3 from (28), getting

$$n = 2^\alpha m = 2^{\alpha'} m' + 2^{\alpha''} m''; \qquad \alpha > 0;$$

$$m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(VII)} \quad \begin{aligned} &\sum [(-1)^{(d'+d'')/2} \{f(d' - d'') + f(d' + d'')\}] \\ &\quad = [\zeta'(n) - \tfrac{1}{2}\mu_1(n)] f(0) \\ &\quad - \sum [f(2) - f(4) + f(6) - \cdots + (-1)^{(d+1)/2} f(d - 1)] \\ &\quad + \sum d[-f(2d) + 2f(2^2 d) + \cdots + 2^{\alpha-1} f(2^\alpha d)]; \end{aligned}$$

and the related form

$$m = 2^{\alpha'} m' + m''; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(VIII)}\quad \begin{aligned} &2\sum [(-1)^{(d'+d'')/2}\{f(d'-d'')+f(d'+d'')\}] \\ &\quad = [\zeta(m)-\zeta_1(m)]f(0) \\ &\qquad -2\sum[f(2)-f(4)+\cdots+(-1)^{(d+1)/2}f(d-1)]. \end{aligned}$$

For $j = 9$ we have

$$2^{\alpha} m = 2^{\alpha'} m' + 2^{\alpha''} m''; \qquad \alpha > 0;$$

$$m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(IX)}\quad \begin{aligned} &\sum [(-1)^{(\delta'+\delta'')/2}\{f(2^{\alpha'}d'-2^{\alpha''}d'')+f(2^{\alpha'}d'+2^{\alpha''}d'')\}] \\ &\quad = [\zeta_1''(2^{\alpha-1}m)-\zeta_1'(2^{\alpha-1}m)]f(0) \\ &\qquad +\sum[(-1)^{(\delta-1)/2}-2^{\alpha}d]f(2^{\alpha}d). \end{aligned}$$

In this the $f(x)$ which appears upon first paraphrasing has been replaced, as allowable, by $f(x/2)$.

$$m = 2^{\alpha'} m' + m''; \qquad m, m', m'' = d\delta, d'\delta', d''\delta''; \qquad \alpha' > 0:$$

$$\text{(X)}\quad \begin{aligned} &2\sum [(-1)^{(\delta'+\delta'')/2}\{f(2^{\alpha'}d'-d'')+f(2^{\alpha'}d'+d'')\}] \\ &\quad = \sum[(-1)^{(\delta-1)/2}-d]f(d). \end{aligned}$$

For $j = 11$, using w_0 from (29), we find

$$n = n' + n''; \qquad n, n', n'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(XI)}\quad \begin{aligned} &\sum [(-1)^{\delta'+\delta''}\{f(d'-d'')-f(d'+d'')\}] \\ &\quad = [\rho(n)-\rho_1(n)]f(0)-\sum[d+(-1)^{\delta}]f(d) \\ &\qquad -2\sum(-1)^{\delta}[f(1)+f(2)+\cdots+f(d-1)]. \end{aligned}$$

The special case in which $n = m$ is of interest:

$$\text{(XII)}\quad \begin{aligned} &\sum [(-1)^{\delta'+\delta''}\{f(d'-d'')-f(d'+d'')\}] \\ &\quad = [\zeta(m)-\zeta_1(m)]f(0)-\sum(d-1)f(d) \\ &\qquad +2\sum[f(1)+f(2)+\cdots+f(d-1)]. \end{aligned}$$

The similarly derived paraphrases of the derivatives $\phi_1' = \phi_7\phi_9$, etc., and those for the relations between pairs of functions whose products are constants, also yield simple and interesting results, but to keep this paper within reasonable limits we pass on to a very brief consideration of the more important series (31)–(46).

16. Let us first paraphrase one of the obvious identities suggested by the form of $\phi_8(x, y)$ in (43), as the resulting paraphrase is one of those which

Pepin ((1), 84–92) proved by Dirichlet's method. By the addition theorems for the ϑ-functions (Jacobi, *Werke*, vol. 1, p. 510, (C)), we have

$$\vartheta_0^2\,\vartheta_1(x+y)\,\vartheta_1(x-y) = \vartheta_1^2(x)\,\vartheta_0^2(y) - \vartheta_0^2(x)\,\vartheta_1^2(y);$$

and hence by (43), (18), on multiplying this identity throughout by

$$\vartheta_2^2\,\vartheta_3^2/\vartheta_1^2(x)\,\vartheta_1^2(x),$$

we get

$$-\phi_8(x,y)\,\phi_8(x,-y) = \phi_2^2(y) - \phi_2^2(x).$$

Substituting in this the respective series as given by (43), (24), replacing q by $\sqrt{q}$, and equating coefficients of q^n, we find

$$n = n' + n''; \qquad n, n', n'' = d\delta, d'\delta', d''\delta'':$$

$$\begin{aligned}\sum[\cot x \sin 2dx \cos 2\delta y - \cot y \cos 2dx \sin 2\delta y]& \\ + \sum[\cos 2\{(d'-d'')x + (\delta'+\delta'')y\}& \\ - \cos 2\{(d'+d'')x + (\delta'-\delta'')y\}] &= \sum d[\cos 2dy - \cos 2dx].\end{aligned}$$

By (29) the first sum*

$$\equiv \sum[\cos 2dy - \cos 2dx] + 2\sum\left[\sum_{r=1}^{d-1}\{\cos 2rx \cos 2\delta y - \cos 2\delta x \cos 2ry\}\right].$$

Considering the second sum (on the left), we have†

$$\begin{aligned}\sum \sin 2(d'-d'')x \sin 2(\delta'+\delta'')y &\equiv 0 \\ &\equiv \sum \sin 2(d'+d'')x \sin 2(\delta'-\delta'')y,\end{aligned}$$

the d' being identical in reversed order with the d'', and similarly for the δ', δ''; hence the second sum reduces to

$$\sum[\cos 2(d'-d'')x \cos 2(\delta'+\delta'')y - \cos 2(d'+d'')x \cos 2(\delta'-\delta'')y].$$

Making all these reductions in the original cot, sin, cos identity, replacing then x, y by $x/2$, $y/2$ and paraphrasing, we get

* d, δ may clearly be interchanged in either term under the Σ:

$$\Sigma[w_0(2d,x)\cos 2\delta y - w_0(2\delta,y)\cos 2dx] \equiv \Sigma[w_0(2d,x)\cos 2\delta y - w_0(2d,y)\cos 2\delta x].$$

We have made this change before writing out the Σ by (29) in the next step; henceforth it will be unnecessary to point out similar transformations.

† The reduction in this step may be obviated by paraphrasing the right of

$$-4\phi_8(x,y)\,\phi_8(x,-y) \equiv [\phi_8(x,y) - \phi_8(x,-y)]^2 - [\phi_8(x,y) + \phi_8(x,-y)]^2,$$

instead of, as above, the left. Such devices sometimes avoid complicated arithmetical reductions.

$$n = n' + n''; \qquad n, n', n'' = d\delta, d'\delta', d''\delta'';$$

$$f(x, y) \equiv f(x, y|):$$

$$\begin{aligned} &\sum [f(d' - d'', \delta' + \delta'') - f(d' + d'', \delta' - \delta'')] \\ &\qquad = \sum [(d-1)\{f(0, d) - f(d, 0)\}] \\ &\qquad\quad + 2\sum \left[\sum_{r=1}^{d-1} \{f(\delta, r) - f(r, \delta)\} \right]. \end{aligned} \tag{XIII}$$

Putting $f(x, y) = f(x)$ or $f(y)$ in (XIII), we find at once

$$n = n' + n''; \qquad n, n', n'' = d\delta, d'\delta', d''\delta'':$$

$$\begin{aligned} &\sum [f(d' - d'') - f(d' + d'')] \\ &\qquad = [\zeta_1(n) - \zeta(n)] f(0) + \sum (2\delta - d - 1) f(d) \\ &\qquad\quad - 2\sum [f(1) + f(2) + \cdots + f(d - 1]). \end{aligned} \tag{XIV}$$

Henceforth we shall write $f(x, y|) \equiv f(x, y)$.

17. By a simple transformation (XIII), (XIV) take more elegant forms. We remark, however, that although slightly simpler in appearance, the new forms are in reality not so simple, containing redundant terms; we give the transformation merely to show the identity of the forms above with Liouville's.

For $n = d\delta$, let $\sum' [\sum_{r=1}^{d-1} f(\delta, r)]$ denote the result of deleting from $\sum [\sum_{r=1}^{d-1} f(\delta, r)]$ every $f(\delta, r)$ for which r is a divisor of d, and similarly for $\sum' [\sum_{r=1}^{d-1} f(r, \delta)]$, $\sum' [\sum_{r=1}^{d-1} f(r)]$. Then it is easily seen that

$$\sum \left[\sum_{r=1}^{d-1} \{f(\delta, r) - f(r, \delta)\} \right] = \sum' \left[\sum_{r=1}^{d-1} f(\delta, r) \right] - \sum' \left[\sum_{r=1}^{d-1} f(r, \delta) \right],$$

$$\begin{aligned} &\sum' [f(2) + f(3) + \cdots + f(d - 1)] \\ &\qquad = \sum [f(1) + f(2) + \cdots + f(d - 1)] - \sum [\zeta(\delta) - 1] f(d). \end{aligned}$$

Hence (XIII), (XIV) over the same separations may be written,

$$\begin{aligned} &\sum [f(d' - d'', \delta' + \delta'') - f(d' + d'', \delta' - \delta'')] \\ &\qquad = \sum (d-1) [f(0, d) - f(d, 0)] \\ &\qquad\quad + 2\sum' \left[\sum_{r=1}^{d-1} \{f(\delta, r) - f(r, \delta)\} \right]; \end{aligned} \tag{XIII'}$$

$$\begin{aligned} \sum [f(d' - d'') - f(d' + d'')] &= [\zeta_1(n) - \zeta(n)] f(0) \\ &\quad - \sum [2\zeta(\delta) + d - 2\delta - 1] f(d) \\ &\quad - 2\sum' [f(2) + f(3) + \cdots + f(d - 1)]; \end{aligned} \tag{XIV'}$$

which are respectively Liouville's (5), 284, (f) and (4), 247, (H).

18. To generalize a paraphrase that arose from ϕ_j^2, we seek $\phi_k(x, y)$, $\phi_l(x, y)$ such that

$$\pm\, \phi_k(x, y)\phi_l(x, -y) = \phi_j^2(x) \pm \phi_j^2(y).$$

Thus for $j = 7$, we have the identity (from (31), and the application of the transformations indicated in the first triads of (21) upon (17))

$$\phi_1(x, y)\phi_1(x, -y) = \phi_7^2(x) - \phi_7^2(y),$$

where we have used

$$\vartheta_3^2\, \vartheta_1(x + y)\,\vartheta_1(x - y) = \vartheta_0^2(x)\,\vartheta_2^2(y) - \vartheta_2^2(x)\,\vartheta_0^2(y).$$

Paraphrasing the ϕ-identity as in § 16, we find immediately

$$2n = m' + m''; \qquad n = 2^\alpha m; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'':$$

$$\text{(XV)}\quad \sum [f(d' - d'', \delta' + \delta'') - f(d' + d'', \delta' - \delta'')] \\ = 2^\alpha \sum d\,[f(0, 2^{\alpha+1} d) - f(2^{\alpha+1} d, 0)],$$

which is Liouville (2), 199, (b), (c), and which becomes (I) for

$$f(x, y) = f(x), f(y).$$

Since by (33), $\phi_3(x, y) = -\phi_1(x', y')$, the paraphrase of

$$\phi_3(x, y)\phi_3(x, -y) \equiv \phi_5^2(x) - \phi_5^2(y)$$

is (XV). From (32), (21) we get

$$\phi_2(x, y)\phi_2(x, -y) = \phi_7^2(x) - \phi_5^2(y),$$

whose paraphrase may be written down from (XV) by means of the elementary transformations of Part I, § 30, on observing that $\phi_2(x, y) = \phi_1(x, y')$. Over the same separation as (XV) it is

$$\text{(XVI)}\quad \sum [(-1)^{(\delta'-1)/2+(\delta''-1)/2}\{f(d' + d'', \delta' - \delta'') + f(d' - d'', \delta' + \delta'')\}] \\ = 2^\alpha \sum d\,[f(2^{\alpha+1} d, 0) - (-1)^n f(0, 2^{\alpha+1} d)].$$

On putting $f(x, y) = f(x)$ in this it becomes (VI), second form; for $f(x, y) = f(y)$ it is the first form of (VI). Continuing thus with the obvious consequences of the developments in § 11, we find from

$$\vartheta_3^2\, \vartheta_0(x + y)\,\vartheta_0(x - y) = \vartheta_0^2(x)\,\vartheta_3^2(y) + \vartheta_2^2(x)\,\vartheta_1^2(y),$$

$$-\phi_4(x, y)\phi_4(x, -y) = \phi_6^2(y) + \phi_7^2(x),$$

on using $u_1(\delta, y)$ from (27), the following paraphrase:

$$
\begin{gathered}
n = n' + n''; \qquad n, n', n'' = 2^{\alpha} m, 2^{\alpha'} m', 2^{\alpha''} m''; \\
m, m', m'' = d\delta, d'\delta', d''\delta'': \\
\begin{aligned}
&\sum [f(2^{\alpha'} d' - 2^{\alpha''} d'', \delta' + \delta'') - f(2^{\alpha'} d' + 2^{\alpha''} d'', \delta' - \delta'')] \\
&\quad = \tfrac{1}{2}\{1 + (-1)^{n}\} \sum [df(0, 2d) + 2df(0, 2^{2} d) + \cdots \\
&\qquad\qquad + 2^{\alpha-1} df(0, 2^{\alpha} d)] - \sum (2^{\alpha} d - 1) f(2^{\alpha} d, 0) \\
&\quad + 2\sum [f(2^{\alpha} d, 2) + f(2^{\alpha} d, 4) + f(2^{\alpha} d, 6) + \cdots \\
&\qquad\qquad + f(2^{\alpha} d, \delta - 1)].
\end{aligned}
\end{gathered}
\tag{XVII}
$$

For $f(x, y) = f(y)$ this becomes (II), (III); for $f(x, y) = f(x)$ we find after a simple reduction,

$$
\begin{gathered}
2^{\alpha} m = 2^{\alpha'} m' + 2^{\alpha''} m''; \qquad m, m', m'' = d\delta, d'\delta', d''\delta''; \\
\alpha > 0: \\
\begin{aligned}
&\sum [f(2^{\alpha'} d' - 2^{\alpha''} d'') - f(2^{\alpha'} d' + 2^{\alpha''} d'')] \\
&\qquad\qquad = \sum (\delta - 2^{\alpha} d)[f(2^{\alpha} d) - f(0)];
\end{aligned}
\end{gathered}
\tag{XVIII}
$$

and when $\alpha = 0$,

$$
\begin{gathered}
m = 2^{\alpha'} m' + m''; \qquad m, m', m'' = d\delta, d'\delta', d''\delta'': \\
2\sum [f(2^{\alpha'} d' - d'') - f(2^{\alpha'} d' + d'')] = \sum (\delta - d) f(d).
\end{gathered}
\tag{XIX}
$$

These three are Liouville's (5), 280, (e); (3), 208, (F); (4), 242, (G). Similarly from (36) we have

$$\phi_5(x, y)\phi_5(x, -y) = \phi_8^2(y) + \phi_7^2(x),$$

which gives upon reduction by $v_3(\delta, y)$ from (28),

$$
\begin{gathered}
n = n' + n''; \qquad n, n', n'' = 2^{\alpha} m, 2^{\alpha'} m', 2^{\alpha''} m''; \\
m, m', m'' = d\delta, d'\delta', d'\delta'': \\
\begin{aligned}
&\sum [(-1)^{(\delta'-1)/2 + (\delta''-1)/2} \{f(2^{\alpha'} d' + 2^{\alpha''} d'', \delta' - \delta'') \\
&\qquad\qquad + f(2^{\alpha'} d' - 2^{\alpha''} d'', \delta' + \delta'')\}] \\
&\quad = \sum (2^{\alpha} d - 1) f(2^{\alpha} d, 0) + 2\sum [f(2^{\alpha} d, 2) - f(2^{\alpha} d, 4) \\
&\qquad\qquad + \cdots + (-1)^{(\delta+1)/2} f(2^{\alpha} d, \delta - 1)] \\
&\quad - \tfrac{1}{2}\{1 + (-1)^{n}\} \sum d[-f(0, 2d) + 2f(0, 2^{2} d) \\
&\qquad\qquad + \cdots + 2^{\alpha-1} f(0, 2^{\alpha} d)],
\end{aligned}
\end{gathered}
\tag{XX}
$$

which might have been derived by the elementary transformations from (XVII); as it is the transformation affords a check of a kind which frequently is valuable. For $f(x, y) = f(x)$ this is easily seen to be

$$
\text{(XXI)}\quad
\begin{aligned}
&\sum [(-1)^{(\delta'-1)/2+(\delta''-1)/2}\{f(2^{\alpha'} d' + 2^{\alpha''} d'') + f(2^{\alpha'} d' - 2^{\alpha''} d'')\}] \\
&\qquad = \tfrac{1}{2}\{1 + (-1)^{n}\}(3 - 2^{\alpha})\zeta_1'(n) f(0) \\
&\qquad\quad + \sum \{2^{\alpha} d + (-1)^{(\delta+1)/2}\} f(2^{\alpha} d),
\end{aligned}
$$

over the same separation as the preceding. This is essentially two theorems; the first corresponds to $\alpha = 0$, and is (X); the second is, with $\alpha > 0$,

$$
\text{(XXII)}\quad
\begin{aligned}
&\sum [(-1)^{(\delta'-1)/2+(\delta''-1)/2}\{f(2^{\alpha'} d' + 2^{\alpha''} d'') + f(2^{\alpha'} d' - 2^{\alpha''} d'')\}] \\
&\qquad = (3 - 2^{\alpha})\zeta_1'(n) + \sum \{2^{\alpha} d + (-1)^{(\delta+1)/2}\} f(2^{\alpha} d).
\end{aligned}
$$

This is Liouville (5), 282, (K). The examples which we have given are probably sufficient as illustrations of the simplest methods of using the tables, and the length of this paper forbids more systematic treatment here of the lists sampled or of the numerous other sets in the theory of elliptic functions which Liouville apparently did not touch. Also we must leave aside the many interesting questions that suggest themselves; e.g., the classification of the paraphrases, the inverse problem of finding paraphrases for a given separation (in which the theory of the transformation of elliptic functions finds a new application to arithmetic), and the passage from paraphrases for functions of degree 2 to degree exceeding 2. One natural starting point for the last is Jacobi's (and H. J. S. Smith's) formula for the multiplication of four theta functions, and its consequences for $\vartheta_a(x \pm y \pm z)$, etc. This generalization is of importance, as it leads naturally to the general case of such paraphrases as the present, viz., to the paraphrase of Riemann's theta formula and the theory of the related functions, making possible a ready arithmetical interpretation of some of the most striking analytical results in the theory of abelian functions.*

In the present order of ideas the series for the $\phi_j^n(x)$, $n > 2$ furnish interesting results; e.g., $\phi_1^3(x)$ gives the principal theorems in Liouville's sixth memoir. We may conclude with two paraphrases for quadratic separations, to illustrate the fact that there is no increase in difficulty by these methods when we pass from linear separations to quadratic. This certainly is not the case in methods used hitherto.

19. From the identity

$$\frac{\vartheta_1'(x)}{\vartheta_1(x)} \times \vartheta_1(x) = \vartheta_1'(x),$$

and the following, which is easily derived from the known series,

* This generalization was carried out in some detail for the hyperelliptic functions of the first order, and in particular for the transformation theory of such functions, in 1915; the results, which lead to interesting arithmetical conclusions, will be published as soon as the papers on the elliptic case have appeared. There is a notable gain in generality when we pass beyond the elliptic case: the majority of the paraphrases refer to functions of n variables unrestricted in any way whatever; parity no longer plays an essential part.

$$\vartheta_1'(x)/\vartheta_1(x) = \cot x + 4\sum q^{2n}\left(\sum \sin 2dx\right),\ [\,T_3\,],$$

we find upon replacing q by q^4, writing $\epsilon(n) = 1$ or 0 according as n is or is not the square of an integer > 0, and paraphrasing as usual;

$$m = 8n' + m''^2; \qquad n' = \delta' d':$$

$$\text{(XXIII)}\quad \begin{aligned} &2\sum(-1)^{(m''-1)/2}[f(2d' - m'') - f(2d' + m'')] \\ &= \epsilon(m)(-1)^{(\sqrt{m}-1)/2}[(\sqrt{m} - 1)f(\sqrt{m}) - 2f(1) - 2f(3) \\ &\qquad - \cdots - 2f(\sqrt{m} - 2)]. \end{aligned}$$

Liouville gave several of the same kind, but not this. He does not seem to have used the $\vartheta_a'(x)/\vartheta_a(x)$.

20. One of the simplest methods for finding quadratic paraphrases concerning functions of order > 1 is by proceeding from identities between theta functions (not their quotients) and series whose d, δ forms are known. Thus it is seen by inspection of (35) that

$$\vartheta_0(x - y)\phi_4(x - y, y) + \vartheta_0(x + y)\phi_4(x + y, -y) \equiv 0,$$

where we have not gone beyond two variables in order to keep the writing simple. Substituting the series for $\vartheta_0(x \pm y)$, $\phi_4(x \mp y, \pm y)$, and noting that $\vartheta_0(x)$ may also be written $\Sigma(-1)^{n_1} q^{n_1^2} \cos 2n_1 x$, where the range of n_1 is $-\infty$ to $+\infty$, we find after some simple reductions the following form of the ϑ, ϕ identity:

$$\begin{aligned} &\sum_{n=1}^{\infty} q^{n^2}\left[(-1)^n \sin 2nx \sum_{r=1}^{n-1} \cos(2r - 1)y\right] \\ &\quad + \sum_{n_1=-\infty}^{\infty} q^{n_1^2+n_2}\left[\sum(-1)^{n_1} \sin(2^{a_2+1} d_2+2n_1)x \cos(2^{a_2+1} d_2-\delta_2+2n_1)y\right] \equiv 0; \end{aligned}$$

and this being an identity in q, we have the next an identity in x, y, (the separation in both cases is the same);

$$n = n_1^2 + n_2; \qquad n_1 \gtreqless 0; \qquad 0 < n_2 = 2^{a_2} m_2; \qquad m_2 = d_2\,\delta_2:$$

$$\begin{aligned} &\epsilon(n)(-1)^n \sin 2\sqrt{n}\,x \sum_{r=1}^{\sqrt{n}} \cos(2r - 1)y \\ &\qquad + \sum(-1)^{n_1} \sin(2^{a_2+1} d_2 + 2n_1)x \cos(2^{a_2+1} d_2 - \delta_2 + 2n_1)y = 0; \end{aligned}$$

and thence, over the same separation,

$$\text{(XXIV)}\quad \begin{aligned} &\sum(-1)^{n_1} f(2^{a_2+1} d_2 - \delta_2 + 2n_1 \,|\, 2^{a_2} d_2 + n_1) \\ &\qquad = \epsilon(n)(-1)^{n-1} \sum_{r=1}^{\sqrt{n}} f(2r - 1 \,|\, \sqrt{n}). \end{aligned}$$

University of Washington

ON THE ZEROS OF SOLUTIONS OF HOMOGENEOUS LINEAR DIFFERENTIAL EQUATIONS[*]

BY

CLARENCE N. REYNOLDS, JR.

1. Introduction

In this paper I shall first prove a general separation theorem for the zeros of solutions of the general homogeneous linear differential equation. I shall then generalize Birkhoff's theorems of oscillation and comparison for equations of the third order[†] by proving two series of theorems, one for equations of odd order and one for equations of even order. This latter series of theorems is then applied to the study of the zeros of the solutions of self-adjoint linear homogeneous differential equations of the fourth order. It would be sufficient for my purposes if I were to assume that the coefficients and solutions of the equations considered, together with a sufficient number of their derivatives, were defined and continuous for all values of the independent variable considered. In order to simplify the statements of my results I shall assume these coefficients and solutions analytic.

We may, without loss of generality, take our nth order equation in the form

$$y^{(n)} + \sum_{i=2}^{n} p_i\, y^{(n-i)} = 0 \qquad (a \leqq x \leqq b) \tag{1}$$

and assume that the wronskian of any fundamental system of those solutions of (1) whose zeros we are studying is identically equal to one; i.e.,

$$W(x) \equiv \begin{vmatrix} y_1, & y_2, & \cdots, & y_n \\ y_1', & y_2', & \cdots, & y_n' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)}, & y_2^{(n-1)}, & \cdots, & y_n^{(n-1)} \end{vmatrix} \equiv 1 .$$

The equation adjoint to (1) is

$$z^{(n)} + \sum_{i=2}^{n} (-1)^i (p_i z)^{(n-i)} = 0 . \tag{2}$$

[*] Presented to the Society, September, 1918, and September, 1919.

[†] Birkhoff, *On the solutions of ordinary linear homogeneous differential equations of the third order*, Annals of Mathematics, ser. 2, vol. 12 (1911), pp. 103–127.

This equation is satisfied by $z_i(x)$, the cofactor of $y_i^{(n-1)}(x)$ in $W(x)$, $i = 1, 2, \cdots, n$. By the properties of $z_i(x)$ as a cofactor,

$$\sum_{i=1}^{n} y_i^{(k)}(x) z_i(x) \equiv 0 \qquad (k = 0, 1, \cdots, n-2), \tag{3}$$

$$\sum_{i=1}^{n} y_i^{(n-1)}(x) z_i(x) \equiv 1. \tag{4}$$

Throughout this paper I shall denote determinants of the form

$$\begin{vmatrix} y_1^{(k_1)}, & y_2^{(k_1)}, & \cdots, & y_n^{(k_1)} \\ y_1^{(k_2)}, & y_2^{(k_2)}, & \cdots, & y_n^{(k_2)} \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(k_n)}, & y_2^{(k_n)}, & \cdots, & y_n^{(k_n)} \end{vmatrix}$$

by the symbol:

$$(y_1^{(k_1)}, y_2^{(k_2)}, \cdots, y_n^{(k_n)}).$$

2. A general separation theorem

THEOREM 1. *If x_1 and x_2 are consecutive zeros of $y_1(x)$, and if $y_2(x)$ does not vanish for either of these values of x, then the number of zeros of $y_2(x)$ between x_1 and x_2 plus the number of zeros of (y_1, y_2') in the same interval is odd.*

If for definiteness we let

$$y_1(x) > 0 \qquad (x_1 < x < x_2),$$

then

$$y_1(x_1) = y_1'(x_1) = \cdots = y_1^{(j-1)}(x_1) = 0, \qquad y_1^{(j)}(x_1) > 0,$$

$$y_1(x_2) = y_1'(x_2) = \cdots = y_1^{(k-1)}(x_2) = 0, \qquad (-1)^{(k)} y_1^{(k)}(x_2) > 0,$$

where $j < n$ and $k < n$ since $W(x) \neq 0$. Differentiating the function $f(x) \equiv y_2(y_1, y_2')$ repeatedly, we find that

$$f^{(m)}(x_1) = 0, \qquad m < j-1,$$

$$f^{(m)}(x_2) = 0, \qquad m < k-1,$$

$$f^{(j-1)}(x_1) = -y_2^2(x_1) y_1^{(j)}(x_1) < 0,$$

$$(-1)^k f^{(k-1)}(x_2) = -(-1)^k y_2^2(x_2) y_1^{(k)}(x_2) < 0,$$

or

$$(-1)^{k-1} f^{(k-1)}(x_2) > 0.$$

Hence the number of zeros of $f(x)$ between x_1 and x_2 is odd. Therefore the number of zeros of $y_1(x)$ plus the number of zeros of (y_1, y_2') between x_1 and x_2 is odd.

Theorem 1 may be geometrically interpreted by supposing the numbers $y_1(x), y_2(x), \cdots, y_n(x)$ to be the homogeneous coördinates of a point on an analytic curve in space of $n-1$ dimensions. This curve is the integral curve of equation (1). The condition $W(x) \neq 0$ then means that the osculating $(n-2)$-way plane spread is never stationary. The determinants (y_i, y'_j) $(i, j = 1, 2, \cdots, n)$, $n(n-1)/2$ in number, are the homogeneous line coördinates of the tangent to the curve at the point $(y_1, y_2, \cdots, y_n)$. The vanishing of (y_1, y'_2) at a point on the curve means that the tangent at the point meets the $(n-3)$-way plane intersection of the $(n-2)$-way plane spreads, whose equations are $y_1 = 0$ and $y_2 = 0$.

Hence Theorem 1 may be read as follows:

If the integral curve of equation (1) does not meet the $(n-3)$-way plane spread whose equations are $y_1 = 0$, $y_2 = 0$, then between two intersections of the curve with the $(n-2)$-way plane spread whose equation is $y_1 = 0$, there are an odd number of intersections with the $(n-2)$-way plane spread whose equation is $y_2 = 0$, and points of tangency with elements of the pencil of $(n-2)$-way plane spreads whose equation is $\alpha y_1 + \beta y_2 = 0$, where α and β are parameters.

For equations of the second order this theorem is equivalent to Sturm's theorem* that the zeros of two linearly independent solutions alternate, since $(y_1, y'_2) \neq 0$ in this case. For equations of the third order it is equivalent to Birkhoff's general separation theorem.†

3. Regular intervals

Let

$$\phi(x, \xi) \equiv \sum_{i=1}^{n} z_i(\xi) y_i(x), \tag{5}$$

then from (3) and (4) we have

$$\frac{\partial^k}{\partial x^k}\phi(x, \xi) = 0, \qquad x = \xi \qquad (k = 0, 1, 2, \cdots, n-2), \tag{6}$$

and

$$\frac{\partial^{n-1}}{\partial x^{n-1}}\phi(x, \xi) = 1, \qquad x = \xi. \tag{7}$$

From (6) and (7) we can prove by substitution that if $\eta(x)$ is a function satisfying

$$\eta(x) \equiv y(x) - \int_a^x \phi(x, \xi) R(\xi) \eta(\xi) d\xi, \tag{8}$$

where $y(x)$ is an arbitrary solution of (1) and $R(x)$ is an arbitrary analytic function of x, then $\eta(x)$ is a solution of the linear differential system

* Cf. C. Sturm, Journal de mathématiques pures et appliquées, vol. 1 (1836), p. 131.

† Cf. Birkhoff, loc. cit., p. 109.

$$\eta^{(n)} + \sum_{i=2}^{n} p_i \, \eta^{(n-i)} + R\eta = 0, \tag{9}$$

$$\eta^{(k)}(\alpha) = y^{(k)}(\alpha) \qquad (k = 0, 1, \cdots, n-1). \tag{10}$$

Interchanging the roles of (1) and (9) we may define

$$\psi(x, \xi) \equiv \sum_{i=2}^{n} \zeta_i(\xi)\, \eta_i(x), \tag{5'}$$

where $\eta_i(x)$, $(i = 1, 2, \cdots, n)$, is a fundamental system of solutions of (9) and the ζ's are derived from the η's as the z's were derived from the y's.

Then

$$\frac{\partial^k}{\partial x^k}\psi(x, \xi) = 0, \qquad x = \xi \quad (k = 0, 1, 2, \cdots, n-2), \tag{6'}$$

$$\frac{\partial^{n-1}}{\partial x^{n-1}}\psi(x, \xi) = 1, \qquad x = \xi, \tag{7'}$$

and

$$y(x) = \eta(x) + \int_a^x \psi(x, \xi) R(\xi) y(\xi) d\xi. \tag{8'}$$

Now, if we set $y(x) = \phi(x, \alpha)$ in (8), we have

$$\psi(x, \alpha) = \phi(x, \alpha) - \int_\xi^x \phi(x, \xi) R(\xi) \psi(\xi, \alpha) d\xi, \tag{11}$$

since $\psi^{(k)}(x, \alpha) = \phi^{(k)}(x, \alpha)$, $x = \xi$, $(k = 0, 1, 2, \cdots, n-1)$, by (6), (7), (6′), and (7′). It will now be possible to compare solutions of (1) and (9) in intervals which satisfy the following

Definition. An interval $(a \leqq x \leqq b)$ shall be said to be regular with respect to (1) and of the first [or second] kind whenever the following conditions are fulfilled, as x and ξ vary throughout the interval:

Kind of Regular Interval	n Odd	n Even
First....................	$\phi(x, \xi) \geqq 0,\ x > \xi$	$\phi(x, \xi) > 0,\ x > \xi$
Second..................	$\phi(x, \xi) \geqq 0,\ x < \xi$	$\phi(x, \xi) < 0,\ x < \xi$.

The proof of the following theorem does not differ essentially from the proof of the special case for the equation of the third order, which Birkhoff has given.* For this reason I do not give the proof here.

THEOREM 2. *If an interval is a regular interval of the first* [*second*] *kind for an equation* (1) *of odd order, then it is also a regular interval of the same kind for the equation* (9), *provided that the inequality* $R(x) \leqq 0\ [R(x) \geqq 0]$ *obtains throughout the interval.*

THEOREM 3. *Any regular interval of either kind for an equation* (1) *of even*

* Loc. cit., p. 117.

order throughout which $R(x) \leqq 0$ is also a regular interval of the same kind for equation (9).

If α is any number in our interval, then, by (6′) and (7′), $\psi(x, \alpha)$ is positive if x is in the immediate right-hand neighborhood of α. If $\psi(x, \alpha) \leqq 0$ for any value of x greater than α, then, since $\psi(x, \alpha)$ is continuous, there exists a number, x_0, greater than α, such that $\psi(x_0, \alpha) = 0, \psi(x, \alpha) > 0$, $(\alpha < x < x_0)$.

If our interval is regular and of the first kind for (1), then $\phi(x_0, \xi) > 0$, $(\alpha \leqq \xi < x_0)$. If we now let $R(x) \leqq 0$ throughout our interval and substitute x_0 for x in (11), we have

$$0 = \phi(x_0, \alpha) - \int_a^{x_0} \phi(x_0, \xi) R(\xi) \psi(\xi, \alpha) d\xi,$$

or zero equal to the sum of a positive quantity and an integral which cannot be negative. Therefore the assumption that $\psi(x, \alpha) \leqq 0$ for any value of x greater than α leads to an absurdity. Hence $\psi(x; \alpha) > 0$ for $x > \alpha$, where α is any value of ξ in our interval, and our interval is regular and of the first kind for the equation (9).

Similarly our theorem may be proven for regular intervals of the second kind.

THEOREM 4. *If $(a \leqq x \leqq b)$ is a regular interval of the first [second] kind, for an equation* (9) *of odd order, throughout which $R(x) < 0$ $[R(x) > 0]$, except for at most a finite number of zeros, and if $y(x)$ and $\eta(x)$ are non-identically vanishing solutions of* (1) *and* (9), *respectively, such that*

$$y^{(k)}(\alpha) = \eta^{(k)}(\alpha), \qquad a \leqq \alpha \leqq b \quad (k = 0, 1, \cdots, n - 1),$$

then between α and the least [greatest] zero of $\eta(x)$ $[y(x)]$ which is greater [less] than α, there exists at least one zero of $y(x)$ $[\eta(x)]$ at which $y(x)$ $[\eta(x)]$ changes sign.

For definiteness, let the first of the numbers $\eta^{(k)}(\alpha)$ which does not vanish be positive. Then, in the immediate right-hand neighborhood of $x = \alpha$, both $\eta(x)$ and $y(x)$ are positive. Now, let $\bar{x}$ be such that $\eta(\bar{x}) = 0$, $\eta(x) > 0$ $(\alpha \leqq x < \bar{x})$. Then, substituting $x = \bar{x}$ in (8′), we have

$$y(\bar{x}) = 0 + \int_a^{\bar{x}} \psi(\bar{x}, \xi) R(\xi) y(\xi) d\xi.$$

Now, if our interval is regular and of the first kind, $\psi(\bar{x}, \xi) \geqq 0 (\bar{x} > \xi)$ and $R(\xi) < 0$ $(\alpha \leqq \xi \leqq \bar{x})$, except for at most a finite number of zeros. Hence, if we suppose that $y(x)$ does not change sign between α and $\bar{x}$, we have $y(\xi) \geqq 0 (\alpha \leqq \xi \leqq \bar{x})$ and a non-negative quantity $y(\bar{x})$ equal to an integral that must be negative, which is absurd. Therefore, $y(x)$ changes sign at least once between α and $\bar{x}$.

If our interval is regular and of the second kind, a similar proof will hold.

If we change the sign of $R(x)$, we interchange $y(x)$ and $\eta(x)$ in the conclusion.

THEOREM 5. *If* $(a \leqq x \leqq b)$ *is a regular interval of the first* [*second*] *kind, for an equation* (9) *of even order, throughout which* $R(x) < 0$, *except for at most a finite number of zeros, and if* $y(x)$ *and* $\eta(x)$ *are non-identically vanishing solutions of* (1) *and* (9), *respectively, such that*

$$y^{(k)}(\alpha) = \eta^{(k)}(\alpha), \qquad (a \leqq \alpha \leqq b) \quad (k = 0, 1, \cdots, n-1),$$

then between α *and the least* [*greatest*] *zero of* $\eta(x)$ [$y(x)$], *which is greater* [*less*] *than* α, *there exists at least one zero of* $y(x)$ [$\eta(x)$] *at which* $y(x)$ [$\eta(x)$] *changes sign.*

A proof similar to that of Theorem 4 proves this theorem, and, as in Theorem 4, a change in sign of $R(x)$ interchanges $y(x)$ and $\eta(x)$ in our conclusion.

4. SELF-ADJOINT EQUATIONS OF THE FOURTH ORDER

The general self-adjoint linear homogeneous differential equation of the fourth order, with the coefficient of the first term identically equal to 1, can be written in the form

$$\eta^{\mathrm{IV}} + 10p_2\,\eta'' + 10p_2'\,\eta' + (3p_2'' + 9p_2^2 + R)\,\eta = 0, \tag{12}$$

where $p_2(x)$ is not the function used in (1). Here the form of the coefficient of η is not dictated by the condition of self-adjointness, but is chosen for the purpose of relating equation (12) to the equation

$$y^{\mathrm{IV}} + 10p_2\,y'' + 10p_2'\,y' + (3p_2'' + 9p_2^2)\,y = 0, \tag{13}$$

which is satisfied by the cube of any solution of*

$$u'' + p_2\,u = 0. \tag{14}$$

Furthermore we know that if $y(x)$ is any solution of (13) then $\eta(x)$ as defined by (8) satisfies (12). If $\phi_2(x, \xi)$ is the solution of (14) defined by (5), then†

$$\phi_2(x, \xi) = 0, \qquad \frac{\partial}{\partial x}\phi_2(x, \xi) = 1, \qquad x = \xi,$$

and

$$\phi_2^3(x, \xi) = \frac{\partial}{\partial x}\phi_2^3(x, \xi) = \frac{\partial^2}{\partial x^2}\phi_2^3(x, \xi) = 0, \qquad \frac{\partial^3}{\partial x^3}\phi_2^3(x, \xi) = 6, \qquad x = \xi.$$

* Cf. Brioschi, Acta Mathematica, vol. 14 (1890), p. 236.

† If $u_1(x)$ and $u_2(x)$ are two solutions of (14) for which $(u_1, u') \equiv 1$, then

$$\phi_2(x, \xi) = -u_2(\xi)\,u_1(x) + u_1(\xi)\,u_2(x).$$

Therefore if $\phi_4(x, \xi)$ is the solution of (13) defined by (5), then

$$\phi_4(x, \xi) = \tfrac{1}{6}\phi_2^3(x, \xi).$$

Hence by the definition of regular intervals any interval which is regular for (14) will be regular for (13) and conversely. Therefore we can substitute (14) for (1) and (12) for (9) in Theorem 3 and obtain

THEOREM 6. *Any regular interval for* (14) *is a regular interval of both kinds for* (12) *provided that* $R(x) \leqq 0$ *throughout our interval.*

Similarly we may substitute (13) for (1), (14) for the first (9) and (12) for the second (9) in Theorem 5 and obtain

THEOREM 7. *If* $(\xi_1 \leqq x \leqq \xi_2)$ *is a regular interval for* (14) *throughout which* $R(x) < 0$ *except for at most a finite number of zeros, and if* $y(x)$ *and* $\eta(x)$ *are non-identically vanishing solutions of* (13) *and* (12) *respectively such that* $y^{(k)}(\alpha) = \eta^{(k)}(\alpha)$, $\xi_1 \leqq \alpha \leqq \xi_2$, $(k = 0, 1, 2, 3)$, *then between* α *and the least* [*greatest*] *zero of* $\eta(x)$ [$y(x)$] *which is greater* [*less*] *than* α *there exists at least one zero of* $y(x)$ [$\eta(x)$] *at which* $y(x)$ [$\eta(x)$] *changes sign.*

In applying these two theorems it is to be noted that sometimes the regular intervals for equation (14) are bounded by two consecutive zeros of a solution of (14), but that if (14) has non-oscillatory solutions then they are not so bounded.

Definitions. The forward interval of oscillation at $x = \alpha$ *for a given equation is the least interval* (α, β) *such that all solutions vanishing for* $x = \alpha$ *will vanish again in* (α, β).

The backward interval of oscillation at $x = \alpha$ *is the least interval* (β, α) *such that all solutions vanishing for* $x = \alpha$ *will vanish again in* (β, α).

THEOREM 8. *If* $R(x) < 0$ *except for at most a finite number of zeros and if equation* (14) *possesses a backward interval of oscillation at* $x = \alpha$, *then equation* (12) *possesses a backward interval of oscillation at* $x = \alpha$ *which is not greater than the backward interval of oscillation for equation* (14).

Any solution of (13), being a homogeneous binary form of the third degree with real constant coefficients in any pair of linearly independent solutions of (14), has at least one real linear factor. This factor vanishes once and only once in the interval $(x_1 \leqq x < \alpha)$ where x_1 is the zero of $\phi_2(x, \alpha)$ which immediately precedes α. Therefore, equation (13) possesses a backward interval of oscillation at $x = \alpha$. Since $\phi_2^3(x, \alpha)$ satisfies (13), this interval is equal to the backward interval of oscillation for (14).

By the preceding theorem any solution, $\eta(x)$, of (12) which vanishes for $x = \alpha$, vanishes and changes sign at least once between α and the greatest zero less than α of that solution, $y(x)$, of (13) which satisfies the conditions (10). Therefore, $\eta(x)$ vanishes and changes sign at least once in the interval $(x_1 \leqq x < \alpha)$. Of all the zeros, infinite in number, immediately preceding α, of the various solutions of (12) which vanish at $x = \alpha$ there must be a lower limiting value, $\bar{x}$, which is not less than x_1, such that all solutions of (12) which vanish for $x = \alpha$ vanish again in the interval $(\bar{x} \leqq x < \alpha)$. This interval is the required interval of oscillation.

5. A separation theorem for self-adjoint equations of the fourth order

If $\eta_i(x)$ and $\eta_j(x)$ satisfy equation (12) then we can prove by differentiating or by Lagrange's identity* that

$$P(\eta_i, \eta_j) \equiv (\eta_i, \eta_j''') - (\eta_i', \eta_j'') + 10p_2(\eta_i, \eta_j') \tag{15}$$

is a constant.

Again it is well known that the six functions

$$(\eta_i, \eta_j') \qquad (i, j = 1, 2, 3, 4, i < j)$$

satisfy a linear differential equation of the fifth order.† Therefore, they must be linearly dependent. If we expand the following linear combination of identically vanishing determinants,

$$(\eta_1, \eta_2''', \eta_3, \eta_4') - (\eta_1', \eta_2'', \eta_3, \eta_4') + 10p_2(\eta_1, \eta_2', \eta_3, \eta_4') \equiv 0, \tag{16}$$

in terms of the two-rowed determinants of the first two rows we have the linear relation

$$\begin{aligned} &P(\eta_1, \eta_2)(\eta_3, \eta_4') - P(\eta_1, \eta_3)(\eta_2, \eta_4') + P(\eta_1, \eta_4)(\eta_2, \eta_3') \\ &\quad + P(\eta_2, \eta_3)(\eta_1, \eta_4') - P(\eta_2, \eta_4)(\eta_1, \eta_3') + P(\eta_3, \eta_4)(\eta_1, \eta_2') \equiv 0. \end{aligned} \tag{17}$$

Furthermore, it can be shown by actual expansion of terms that

$$\zeta_1 \equiv P(\eta_2, \eta_3)\eta_4 + P(\eta_3, \eta_4)\eta_2 + P(\eta_4, \eta_2)\eta_3, \tag{18}$$

the other ζ's being obtained by cyclic permutations of the subscripts. Substituting ζ_1 in (17) we have

$$(\eta_1, \zeta_1') + \begin{vmatrix} \eta_2, & \eta_3, & \eta_4 \\ \eta_2', & \eta_3', & \eta_4' \\ P(\eta_1, \eta_2), & P(\eta_1, \eta_3), & P(\eta_1, \eta_4) \end{vmatrix} \equiv 0, \tag{19}$$

where our three-rowed determinant does not involve $\eta_1(x)$ explicitly. Solving for (η_1, ζ_1') and substituting in Theorem 1, we have a general separation theorem for self-adjoint equations of the fourth order.

THEOREM 9. *If $\eta_i(x)$ $(i = 1, 2, 3, 4)$ are linearly independent solutions of* (12) *then between two consecutive zeros of $\eta_1(x)$ at which*

* Cf. Bôcher, *Leçons sur les Méthodes de Sturm*, Paris, 1917, p. 23.

† Cf. Forsyth, Philosophical Transactions, vol. 179 (1888), p. 455.

$$\zeta_1(x)[\equiv -(\eta_2, \eta_3', \eta_4'')]$$

does not vanish, there exist an odd number of zeros of $\zeta_1(x)$ *and*

$$\begin{vmatrix} \eta_2, & \eta_3, & \eta_4 \\ \eta_2', & \eta_3', & \eta_4' \\ P(\eta_1, \eta_2), & P(\eta_1, \eta_3), & P(\eta_1, \eta_4) \end{vmatrix}.$$

Interpreting this theorem geometrically by means of the curve whose parametric equations are $y_i = \eta_i(x)$ $(i = 1, 2, 3, 4)$ and the conical projection of this curve upon the plane $y_1 = 0$ from the vertex $(1, 0, 0, 0)$ we see that between two consecutive intersections of our curve with the plane $y_1 = 0$ at which the projected curve has no point of inflection there will exist on the projected curve an odd number of points of inflection and points at which the tangent to the curve passes through the point $[0, P(\eta_1, \eta_2), P(\eta_1, \eta_3), P(\eta_1, \eta_4)]$.

6. Applications

In this section I shall apply the theorems derived in Section 4 to two examples.

Example 1.

$$(a) \qquad \eta^{\mathrm{IV}} + 10\eta'' + x\eta = 0, \qquad x \leqq 9.$$

We shall compare the solutions of this equation with the solutions of the equation

$$(b) \qquad y^{\mathrm{IV}} + 10y'' + 9y = 0.$$

Here $R(x) = x - 9 \leqq 0$, and (*b*) is satisfied by $u^3(x)$ whenever

$$(c) \qquad u'' + u = 0.$$

Every interval of length π is an interval of oscillation for equations (*b*) and (*c*). Therefore, by Theorem 6, $(9 - \pi \leqq x \leqq 9)$ is a regular interval of both kinds for (*a*). Theorem 7 implies that if two solutions, $\eta(x)$ and $y(x)$, of (*a*) and (*b*) respectively, are such that

$$\eta^{(k)}(\alpha) = y^{(k)}(\alpha), \qquad 9 - \pi \leqq \alpha \leqq 9 \qquad (k = 0, 1, 2, 3),$$

then between α and the least [greatest] zero of $\eta(y)$ which is greater than α [less than α] there exists at least one zero of $y(\eta)$ at which $y(\eta)$ changes signs. By Theorem 8 we know that for $\alpha \leqq 9$, (*a*) possesses a backward interval of oscillation at $x = \alpha$ of length not greater than π.

Example 2.

$$(a) \qquad \eta^{\mathrm{IV}} - x\eta = 0, \qquad x \geqq 0.$$

Here $R(x) = -x \leqq 0$. In this case our comparison equation

$$(b) \qquad y^{\mathrm{IV}} = 0$$

has no intervals of oscillation. If the particular solution $\eta_1(x)$ satisfies the boundary conditions

$$(c) \qquad \eta_1(\alpha) = 1, \qquad \eta_1'(\alpha) = \eta_1''(\alpha) = \eta_1'''(\alpha) = 0, \qquad \alpha \geqq 0,$$

then $y_1(x)$, the corresponding solution of (*b*), is identically equal to 1. Therefore, by Theorem 7, $\eta_1(x)$ cannot vanish for any value of x greater than α. Similarly none of the other principal solutions of (*a*) at $x = \alpha$ vanish for x greater than α.

If $\eta_2(x)$ satisfies the boundary conditions

$$(c') \quad \eta_2(1) = 0, \qquad \eta_2'(1) = -1, \qquad \eta_2''(1) = 0, \qquad \eta_2'''(1) = 6,$$

then

$$y_2(x) = x(x-1)(x-2).$$

Therefore by Theorem 7, $\eta_2(x)$ has at least one zero in the interval $(0 \leqq x \leqq 1)$, no zero in the interval $(1 \leqq x \leqq 2)$, and may have zeros for x greater than 2.

A GENERALIZATION OF THE FOURIER COSINE SERIES*

BY

J. L. WALSH

It is well known that on the interval $0 \leqq x \leqq \pi$ large classes of functions arbitrary except for certain restrictions as to oscillation or as to continuity and the existence and continuity of derivatives or as to some similar property can be developed into series in terms of the functions $\{\cos nx\}$. The series are of the type

$$(1)\qquad f(x) = \frac{a_0}{2} + a_1 \cos x + a_2 \cos 2x + \cdots,$$

where any coefficient a_n may be obtained by the formal process of multiplying equation (1) through by $\cos nxdx$ and integrating term by term:

$$a_n = \frac{2}{\pi}\int_0^\pi f(x)\cos nxdx.$$

The present paper considers the problem of developing arbitrary functions throughout the same interval in terms of the second set of functions $\{\cos \lambda_n x\}$, where λ_n is very near to n. Under suitable restrictions it is proved that the two developments have essentially the same convergence properties.† We

* Presented to the Society, September, 1920.

† This study was undertaken at the suggestion of Professor Birkhoff, who proposed an analogous generalization of the Fourier sine series. There is no analogous generalization of the sine series such that developments in terms of the two sets of functions $\{\sin nx\}$ and $\{\sin \lambda_n x\}$, where λ_n is near but not equal to n, will have the same convergence properties on the interval $0 \leqq x \leqq \pi$. For at the right-hand end of the interval any particular function $\sin \lambda_k x$ is not zero and hence cannot be developed in a uniformly convergent series in terms of $\{\sin nx\}$, but it can be developed in a uniformly convergent series in terms of $\{\sin \lambda_n x\}$. Thus the methods of the present paper cannot be used to give a generalization of the sine series on the interval $0 \leqq x \leqq \pi$, but these methods with slight modifications would enable us to replace the set $\{\sin nx\}$ by a set $\{\sin \lambda_n x\}$ such that developments in terms of the two sets of functions will have the same convergence properties on the interval $0 \leqq x \leqq \pi - \epsilon$, where ϵ is arbitrary.

The methods used in this paper are in spirit closely related to the methods used by Birkhoff, Paris Comptes Rendus, vol. 164 (1917), pp. 942–945, who gives a similar generalization of Taylor's series. Birkhoff makes use of the theory of integral equations, but the related theory of infinitely many variables as used in the present paper leads to more general results.

shall say that two series

$$\sum_{i=1}^{\infty} a_i \bar{u}_i, \qquad \sum_{i=1}^{\infty} b_i u_i$$

have essentially the same convergence properties when and only when the series

$$\sum_{i=1}^{\infty} (a_i \bar{u}_i - b_i u_i)$$

converges absolutely and uniformly to the sum zero.

The set of functions

$$\frac{1}{\pi}, \ \frac{2}{\pi}\cos x, \ \frac{2}{\pi}\cos 2x, \ \cdots$$

is a normal and orthogonal set on the interval $0 \leqq x \leqq \pi$. Any set of functions $\{\bar{u}_n\}$ continuous on a finite interval is said to be normal and orthogonal on that interval if and only if

$$\int \bar{u}_i \bar{u}_j \, dx = \delta_{ij} \qquad (i, j = 1, 2, 3, \cdots);$$

here and throughout this paper δ_{ij} is the Kronecker symbol which is unity or zero according as i and j are or are not equal. In this integral and below we omit the argument, which is x in every case, and also omit the limits of integration, which are the ends of the interval considered. We shall prove that corresponding to the set of functions

$$\frac{1}{\pi}\cos \lambda_0 x, \ \frac{2}{\pi}\cos \lambda_1 x, \ \frac{2}{\pi}\cos \lambda_2 x, \ \cdots$$

there is another set of functions such that the two sets are biorthogonal for $0 \leqq x \leqq \pi$. Two sets of continuous functions $\{u_n\}$ and $\{v_n\}$ are biorthogonal on an interval if and only if

$$\int u_i v_j \, dx = \delta_{ij} \qquad (i, j = 1, 2, \cdots).$$

We shall prove a general theorem concerning the generalization of expansions in terms of any uniformly bounded normal orthogonal set $\{\bar{u}_n\}$, by considering a set $\{u_n\}$ where u_n is a function *neighboring* to $\bar{u}_n$. Briefly described, the method is to expand the u_n into series in terms of the $\bar{u}_n$, to invert the corresponding system of equations and expand the $\bar{u}_n$ in terms of the u_n, and finally to substitute the latter series into the expansion of the arbitrary function in terms of the $\bar{u}_n$. This gives the expansion of $f(x)$ in terms of the u_n. The form of presentation of the material is slightly changed by the introduction of a set of functions $\{v_n\}$ such that $\{u_n\}$ and $\{v_n\}$ are biorthogonal sets. The various steps in the proof will seem much less artificial if the general method is kept in mind.

After the general theorem is proved, application is made to the cosine series. We add that the method seems to be of very wide applicability; it is hoped later to give further applications to the Sturm-Liouville and other series developments. Throughout the present paper there is more regard for simplicity than for ultimate generality.

I. A general theorem

THEOREM I. *Suppose that* $\{\bar{u}_n\}$ *is a set of uniformly bounded normal orthogonal functions in an interval, and that in this interval* $\{u_n\}$ *is a set of uniformly bounded continuous functions each of which can be developed into a series*

$$u_n = \sum_{k=1}^{\infty} (c_{nk} + \delta_{nk}) \bar{u}_k \qquad (n = 1, 2, \cdots), \tag{2}$$

where the coefficients have the values

$$c_{nk} + \delta_{nk} = \int u_n \bar{u}_k \, dx. \tag{3}$$

Suppose further that the three series

$$\sum_{i,k=1}^{\infty} c_{ik}^2, \quad \sum_{i=1}^{\infty} \left(\sum_{k=1}^{\infty} c_{ik}^2 \right)^{\frac{1}{2}}, \quad \sum_{k=1}^{\infty} \left(\sum_{i=1}^{\infty} c_{ik}^2 \right)^{\frac{1}{2}} \tag{4}$$

converge and that the value of the first is less than unity.

Then there exists a set of functions $\{v_i\}$ *such that* $\{u_i\}$ *and* $\{v_i\}$ *are biorthogonal sets*:

$$\int u_i v_j \, dx = \delta_{ij} \qquad (i, j = 1, 2, \cdots). \tag{5}$$

Furthermore, if $f(x)$ *is any function integrable and with an integrable square (in the sense of Lebesgue), then the two series*

$$f(x) \sim \sum_{i=1}^{\infty} a_i \bar{u}_i, \tag{6}$$

$$f(x) \sim \sum_{i=1}^{\infty} b_i u_i, \tag{7}$$

where

$$\begin{aligned} a_i &= \int f \bar{u}_i \, dx, \\ b_i &= \int f v_i \, dx, \end{aligned} \tag{8}$$

have essentially the same convergence properties.

The sign $\sim$ is used simply to indicate that the coefficients a_i and b_i are given by (8). We shall give later some of the more immediate consequences of the fact that (6) and (7) have essentially the same convergence properties.

To prepare for the inversion of system (2), which is an important step in the proof of Theorem I, we shall use the following

LEMMA: *If for the system*

$$(9)\qquad \begin{array}{rrrl} (1+c_{11})x_1 + & c_{12}\ x_2 + & c_{13}\ x_3 + \cdots & = c_1, \\ c_{21}\ x_1 + & (1+c_{22})x_2 + & c_{23}\ x_3 + \cdots & = c_2, \\ c_{31}\ x_1 + & c_{32}\ x_2 + & (1+c_{33})x_3 + \cdots & = c_3, \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ , \end{array}$$

the series

$$\gamma^2 = \sum_{j=1}^{\infty} c_j^2, \qquad p^2 = \sum_{i,j=1}^{\infty} c_{ij}^2$$

converge, and if $p^2 < 1$, *then the system has one solution* $\{x_i\}$ *and only one solution for which*

$$\sum_{i=1}^{\infty} x_i^2$$

*converges.**

The lemma is proved by the method of successive approximations, setting

$$\begin{aligned} x_i^{(1)} &= c_i, \\ x_i^{(2)} &= c_i - [c_{i1}x_1^{(1)} + c_{i2}x_2^{(1)} + \cdots], \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ , \\ x_i^{(n+1)} &= c_i - [c_{i1}x_1^{(n)} + c_{i2}x_2^{(n)} + \cdots], \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ . \end{aligned}$$

From the well-known inequality

$$(10)\qquad \sum_{k=1}^{n} |\alpha_k \beta_k| \leqq \left(\sum_{k=1}^{n} \alpha_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} \beta_k^2\right)^{\frac{1}{2}}$$

(which holds for all sets of numbers $\{\alpha_k\}$ and $\{\beta_k\}$) and from the convergence of the series whose sums are denoted by γ^2 and p^2, it follows that the series

$$x_i = x_i^{(1)} + (x_i^{(2)} - x_i^{(1)}) + (x_i^{(3)} - x_i^{(2)}) + \cdots \qquad (i = 1, 2, \cdots)$$

are all absolutely convergent and that the system $\{x_i\}$ as thus defined is a

* This lemma is included implicitly in some results of Hilb, Sitzungsberichte der Physikalisch-medizinischen Sozietät in Erlangen, vol. 40 (1908), pp. 1–6. Hilb makes use of the theory of bilinear forms in infinitely many variables, so that his proof differs in form (but not in substance) from the proof here presented. We indicate the proof because of its simplicity and because we shall use later the inequalities obtained in the course of the proof.

solution of (9). If we introduce the notation $p_i^2 = \sum_{j=1}^{\infty} c_{ij}^2$, it is readily proved by further use of the same inequality that

$$|x_i - c_i| \leqq \frac{\gamma p_i}{1-p},\qquad\qquad (11)$$

$$\sum_{i=1}^{\infty} |x_i - c_i|^2 \leqq \frac{\gamma^2 p^2}{(1-p)^2}.$$

The series $\sum_{i=1}^{\infty} x_i^2$ converges, for we obtain the relations*

$$|x_i| \leqq |x_i - c_i| + |c_i|,$$

$$\sum_{i=1}^{\infty} x_i^2 \leqq \sum_{i=1}^{\infty} (|x_i - c_i| + |c_i|)^2 \leqq 2\sum_{i=1}^{\infty} (x_i - c_i)^2 + 2\sum_{i=1}^{\infty} c_i^2.$$

There is no other solution of system (9) such that $\sum_{i=1}^{\infty} x_i^2$ converges. For if there were, we should have a solution $\{y_i\}$ of the homogeneous system corresponding to (9) and such that the series

$$Y^2 = \sum_{i=1}^{\infty} y_i^2$$

converges. In the kth equation of the homogeneous system, if we transpose y_k to the right-hand member, we obtain

$$y_k^2 \leqq Y^2 p_k^2,$$

and hence we also have

$$\sum_{k=1}^{\infty} y_k^2 \leqq Y^2 p^2.$$

In this same manner we have

$$\sum_{k=1}^{\infty} y_k^2 \leqq Y^2 p^{2n},$$

where n is any positive integer, and hence $y_k = 0$ $(k = 1, 2, \cdots)$.

We apply this lemma (the proof of which is now complete) to determine the set of functions $\{v_k\}$ of Theorem I. Let us assume a development of the type

$$v_k = \sum_{i=1}^{\infty} (d_{ki} + \delta_{ki})\bar{u}_i, \qquad \text{where} \qquad d_{ki} + \delta_{ki} = \int v_k \bar{u}_i \, dx. \qquad (12)$$

The condition that the v_k have the desired property (5) is given formally from the developments (2) and (12) by the conditions

$$c_{ki} + d_{ik} + c_{k1} d_{i1} + c_{k2} d_{i2} + c_{k3} d_{i3} + \cdots = 0 \qquad (13)$$

for every i and k. When i is kept fixed, this is precisely system (9); the

* By use of the inequality

$$(\alpha + \beta)^2 \leqq 2(\alpha^2 + \beta^2).$$

conditions of the lemma are fulfilled by virtue of the hypothesis of Theorem I, and therefore there exists a solution $(d_{i1}, d_{i2}, d_{i3}, \cdots)$,* with the properties stated in the lemma. If we choose that solution and use inequalities (11) and the convergence of series (4), we actually arrive at a system of continuous functions $\{v_k\}$ which satisfy (5) and are given by (12) and (13).

Our preparation for the inversion of system (2) is nearly complete. We introduce the notation

$$\epsilon_{ik} = c_{ki} + d_{ik} + c_{1i}\, d_{1k} + c_{2i}\, d_{2k} + c_{3i}\, d_{3k} + \cdots \tag{14}$$

and shall prove $\epsilon_{ik} = 0$ for every i and k. In the set of equations

$$\begin{aligned}
\epsilon_{i1} &= c_{1i} + d_{i1} + c_{1i}\, d_{11} + c_{2i}\, d_{21} + c_{3i}\, d_{31} + \cdots, \\
\epsilon_{i2} &= c_{2i} + d_{i2} + c_{1i}\, d_{12} + c_{2i}\, d_{22} + c_{3i}\, d_{32} + \cdots, \\
\epsilon_{i3} &= c_{3i} + d_{i3} + c_{1i}\, d_{13} + c_{2i}\, d_{23} + c_{3i}\, d_{33} + \cdots, \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot,
\end{aligned}$$

multiply the ith equation by c_{ni} and sum by columns, making use of (13) and (14). We obtain the result

$$c_{k1}\, \epsilon_{i1} + c_{k2}\, \epsilon_{i2} + c_{k3}\, \epsilon_{i3} + \cdots = -\, \epsilon_{ik}.$$

The formal work is readily justified by the convergence of series (4). Moreover, the series $\sum_{k=1}^{\infty} \epsilon_{ik}^2$ converges, from the definition (14) and from (11) applied to the set $(d_{i1}, d_{i2}, d_{i3}, \cdots)$ as a solution of (13), so the set $(\epsilon_{i1}, \epsilon_{i2}, \epsilon_{i3}, \cdots)$ forms a solution of the homogeneous system corresponding to (9) and hence $\epsilon_{ik} = 0$ for every i and k.

We are now in a position to invert system (2). We write the equations

$$\begin{aligned}
c_{k1} + d_{1k} + c_{11}\, d_{1k} + c_{21}\, d_{2k} + c_{31}\, d_{3k} + \cdots &= 0, \\
c_{k2} + d_{2k} + c_{12}\, d_{1k} + c_{22}\, d_{2k} + c_{32}\, d_{3k} + \cdots &= 0, \\
c_{k3} + d_{3k} + c_{13}\, d_{1k} + c_{23}\, d_{2k} + c_{33}\, d_{3k} + \cdots &= 0, \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad &\cdot,
\end{aligned}$$

multiply the ith equation by $\bar{u}_i$ and sum by columns. Justification of the formal work is immediate, and we obtain the desired developments

$$\bar{u}_k - u_k = \sum_{i=1}^{\infty} d_{ik}\, u_i \qquad (k = 1, 2, 3, \cdots).$$

Let us suppose, now, $f(x)$ to be any function integrable and with an integrable square in the sense of Lebesgue. We find the coefficients (8) by multi-

[*] For a non-vanishing finite determinant, the numbers $d_{ik} + \delta_{ik}$ are the quotients of the cofactors of the $c_{ik} + \delta_{ik}$ by the value of the determinant.

plying (12) through by $f(x)$ and integrating term by term. This yields the result

$$(15) \qquad b_k - a_k = d_{k1}\, a_1 + d_{k2}\, a_2 + d_{k3}\, a_3 + \cdots .^{*}$$

The difference of the series (6) and (7) may be written

$$(16) \qquad \begin{aligned} \sum_{1=i}^{\infty} (a_i\, \bar{u}_i - b_i\, u_i) &= \sum_{i=1}^{\infty} a_i (\bar{u}_i - u_i) + \sum_{i=1}^{\infty} (a_i - b_i)\, u_i \\ &= - \sum_{i=1}^{\infty} a_i \left(\sum_{j=1}^{\infty} c_{ij}\, \bar{u}_j \right) - \sum_{i=1}^{\infty} (d_{i1}\, a_1 + d_{i2}\, a_2 + \cdots) \left(\bar{u}_i + \sum_{j=1}^{\infty} c_{ij}\, \bar{u}_j \right), \end{aligned}$$

which series converges absolutely and uniformly and has the sum zero. We make use of the convergence of the series $\sum_{i=1}^{\infty} a_i^2$, which may be proved easily from the relation

$$(17) \qquad \int (f - a_1\, \bar{u}_1 - a_2\, \bar{u}_2 - \cdots - a_n\, \bar{u}_n)^2\, dx \geqq 0 .$$

Hence, *series* (6) *and* (7) *have the same convergence properties, in the sense that* (16) *converges absolutely and uniformly to zero.* It follows immediately that properties of absolute convergence, convergence, summability, and divergence obtain for one series as for the other, and likewise the properties of uniform convergence in the entire interval considered or in any sub-interval, uniform summability, and also of term-by-term integrability. Whenever the two series (6) and (7) are convergent, summable, or properly divergent, their sums are the same. The nature of the approximating functions and of their approach to the limit (in case of convergence) at a point of continuity or of discontinuity of $f(x)$ is essentially the same for (6) as for (7). In particular if Gibbs's phenomenon occurs for (6) it also occurs for (7). The reader will notice various other properties common to the sets $\{\bar{u}_i\}$ and $\{u_i\}$, such as the existence or non-existence of a continuous function for which the formal series does not converge at every point.

The conditions of Theorem I concerned with the series (4) can be made

[*] If we use the convergence of $\Sigma_{i=1}^{\infty} a_i^2$, which is proved from (17); if we write the equations

$$\begin{aligned} b_1 - a_1 &= d_{11}\, a_1 + d_{12}\, a_2 + d_{13}\, a_3 + \cdots , \\ b_2 - a_2 &= d_{21}\, a_1 + d_{22}\, a_2 + d_{23}\, a_3 + \cdots , \\ b_3 - a_3 &= d_{31}\, a_1 + d_{32}\, a_2 + d_{33}\, a_3 + \cdots , \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot , \end{aligned}$$

multiply the ith equation by c_{ik}, and sum by columns, we obtain

$$a_k - b_k = c_{1k}\, b_1 + c_{2k}\, b_2 + c_{3k}\, b_3 + \cdots .$$

Hence if the $\{b_i\}$ are all zero, so are all the $\{a_i\}$. The converse is evident from (15).

We also add that if the series $\Sigma_{i=1}^{\infty} |a_i|$ converges, so does the series $\Sigma_{i=1}^{\infty} |b_i|$, and conversely.

more restrictive so that the convergence properties of (6) and (7) will be essentially the same for all functions absolutely integrable in the sense of Lebesgue, or on the other hand can be made less restrictive and still yield a result say for functions such that the series of absolute values of the coefficients $\sum_{i=1}^{\infty}|a_i|$ converges.

We add the remark that the relation between the $\{u_i\}$ and the $\{v_i\}$ is essentially reciprocal—we may obtain corresponding results for expansions in terms of the $\{v_i\}$. Theorem I is readily extended, moreover, by taking as point of departure developments in terms of functions $\{\bar{u}_i\}$ which are not supposed orthogonal.

II. Application to the Fourier cosine series

For the interval $0 \leqq x \leqq \pi$, we have mentioned the normal orthogonal system of uniformly bounded continuous functions

$$\frac{1}{\pi},\ \frac{2}{\pi}\cos x,\ \frac{2}{\pi}\cos 2x,\ \frac{2}{\pi}\cos 3x,\ \cdots. \tag{18}$$

Application of Theorem I will yield the following theorem.

Theorem II. *The expansion of any function $f(x)$ integrable and with an integrable square in terms of the functions*

$$\frac{1}{\pi}\cos\lambda_0 x,\ \frac{2}{\pi}\cos\lambda_1 x,\ \frac{2}{\pi}\cos\lambda_2 x,\ \frac{2}{\pi}\cos\lambda_3 x,\ \cdots$$

on the interval $0 \leqq x \leqq \pi$, where

$$\lambda_0^2 + 4(\lambda_1 - 1)^2 + 4(\lambda_2 - 2)^2 + 4(\lambda_3 - 3)^2 + \cdots < \frac{1}{\pi} \tag{19}$$

and where

$$\sum_{n=1}^{\infty} n^2|\lambda_n - n| \tag{20}$$

converges, has essentially the same convergence properties as the expansion of $f(x)$ in terms of the functions (18).

It is of course true that our set of functions $\{u_m\}$ can be developed in terms of the functions $\{\bar{u}_m\}$, so it merely remains to prove the convergence of the series (4) and that the value of the first of these series is less than unity.

Inequality (17) gives us the relation

$$\sum_{k=1}^{\infty} c_{mk}^2 \leqq \int(\bar{u}_m - u_m)^2\,dx \qquad (m = 1, 2, 3, \cdots),$$

and in the present case if $m \neq 1$ and if we replace m by $n + 1$, this integral becomes

$$\frac{4}{\pi^2}\int_0^\pi (\cos\lambda_n x - \cos nx)^2\,dx \tag{21}$$
$$= \frac{4}{\pi^2}\int_0^\pi [4\sin^2 \tfrac{1}{2}(\lambda_n + n)x \cdot \sin^2 \tfrac{1}{2}(\lambda_n - n)x]\,dx.$$

But we always have the relation $|\sin\alpha| \leqq |\alpha|$, and hence we obtain

$$\sum_{k=1}^{\infty} c_{n+1,\,k}^2 \leqq 4\pi(\lambda_n - n)^2 \qquad (n = 1, 2, 3, \cdots).$$

If we omit the factor 4 from the right-hand member we obtain the proper formula for $n = 0$. Satisfaction of (19) therefore means that we have satisfied the following requirement of Theorem I

$$\sum_{m,\,k=1}^{\infty} c_{mk}^2 < 1.$$

It remains to be shown that the last two series in (4) converge. From the formula

$$c_{n+1,\,k+1} = \frac{4}{\pi^2}\int_0^\pi (\cos\lambda_n x - \cos nx)\cos kx dx \tag{22}$$
$$(n = 1, 2, \cdots; k = 1, 2, \cdots)$$

integration by parts twice yields the formula

$$\frac{\pi^2}{4}c_{n+1,\,k+1} = -\left[(\lambda_n \sin\lambda_n x - n\sin nx)\frac{\cos kx}{k^2}\right]_0^\pi \tag{23}$$
$$+ \int_0^\pi (\lambda_n^2\cos\lambda_n x - n^2\cos nx)\frac{\cos kx}{k^2}\,dx.$$

To deal with this integral, we write

$$\lambda_n^2\cos\lambda_n x - n^2\cos nx = (\lambda_n^2 - n^2)\cos\lambda_n x + n^2(\cos\lambda_n x - \cos nx). \tag{24}$$

If $n > 2$, we find with the aid of (19) the obvious relation

$$|\lambda_n + n| < 2n + 2 < n^2,$$

whence

$$|\lambda_n^2 - n^2| < n^2|\lambda_n - n|;$$

we use this inequality in considering the first term of the right-hand member of (24). The last term of the right-hand member of (24) can be transformed as in (21), and its absolute value is therefore not greater than $n^2\pi|\lambda_n - n|$.

The first term in the right-hand member of (23) with the factor $\cos k\pi/k^2$ omitted, reduces to $\lambda_n(\sin\lambda_n\pi - \sin n\pi)$. When $n > 1$, we have $|\lambda_n| \leqq n^2$, and therefore

$$|\lambda_n(\sin\lambda_n\pi - \sin n\pi)| \leqq n^2\pi|\lambda_n - n|.$$

From these inequalities for the absolute value of the right-hand member of (23) and from the corresponding results for $k = 0$ and $n \leqq 2$ we prove the convergence of

$$\sum_{n,\,k=1}^{\infty} |c_{nk}|$$

by means of the convergence of (20) and hence* prove the convergence of the last two series of (4) and complete the proof of Theorem II.

We add the remark that a closer evaluation of c_{nk}, and also the consideration of functions $u_n = \mu_n \cos \lambda_n x$ where μ_n is properly determined, will give broader restrictions than (19) and (20).

HARVARD UNIVERSITY,
May, 1920

* Here we make use of the inequality

$$\sum_{i=1}^{n} |\alpha_i| = \left[\left(\sum_{i=1}^{n} |\alpha_i|\right)^2\right]^{\frac{1}{2}} \geqq \left[\sum_{i=1}^{n} \alpha_i^2\right]^{\frac{1}{2}}.$$

POLYNOMIALS AND THEIR RESIDUE SYSTEMS*

BY

AUBREY J. KEMPNER

Introduction

The present paper develops an elementary theory of polynomials (with rational integral coefficients) with respect to a modulus m, where m is a given *composite rational integer,* and of residue systems of such polynomials with respect to a modulus m. Only isolated results have been previously established.

It is known that, for p a prime, and any set of integers α_i $(i = 0, 1, \cdots, p-1)$, there exist polynomials with integral coefficients of which the set form a complete residue system* modulo p such that $f(i) \equiv \alpha_i$. The simplest considerations show that this is not true for a composite modulus, m. In this case certain relations between the α_i $(i = 0, 1, \cdots, m-1)$ must be satisfied if they are to form a complete residue system modulo m of a polynomial with integral coefficients. The problem of establishing necessary and sufficient conditions for the α_i to be, for a given modulus m, such a residue system, and the examination of such systems, was the starting point for the present article. The existence of a certain type of isomorphism between the structure of the totality of reduced polynomials modulo m on one hand and the totality of complete residue systems modulo m on the other hand suggested an independent parallel development of the theory of the polynomials and the theory of th? residue systems. This has been carried out in Part I (Residual congruences and completely reduced polynomials) and Part III (Residual congruences and residue systems).

After the article was written, I became acquainted with several papers on Kronecker modular systems which establish in a satisfactory manner relations between Part I and the well developed theory of modular systems. The short Part II, which was then inserted, is devoted to a brief discussion of this connection.

The classical methods of the theory of numbers may be said to have a

* Presented to the Society, Chicago, December, 1920.

* See, for example, Zsigmondy, Monatshefte für Mathematik und Physik, vol. 8 (1897), p. 20.

tendency to consider an investigation closed when the problem has been reduced to the treatment of a number of cases where the modulus is a prime or a power of a prime, leaving the synthesis of the general case from these special cases in descriptive form. The results and methods of the paper may make it possible to break down, in some fields of investigation, these barriers between the case of a prime modulus or a prime-power modulus, and the case of a general composite modulus.

It is certain that this cannot be accomplished unless the methods are general, lead without trials to clean-cut results, and are readily applicable to a given numerical case. The last demands will justify the inclusion of numerical examples throughout the work. They also made it desirable to treat in detail some of the simpler types (m a prime; m a product of distinct primes; m a power of a prime, with the sub-cases $m = p^\gamma, \gamma < p$; $m = p^\gamma, \gamma \geqq p$). This involves only a small increase in space, since the proof of the formulæ for the general case is naturally based on the simpler cases.

The methods employed tend to emphasize, for a given modulus, the totality of completely reduced polynomials modulo m rather than the individual polynomial, and the totality of complete residue systems modulo m rather than the individual residue system.*

Part III, however, also furnishes tools for the examination of the individual residue system. Some applications of the theory are reserved for another occasion.

The introduction of some new terms and symbols could not well be avoided. A brief synopsis may be of assistance to the reader:

In Part I, for a given modulus, m, *a certain finite set of polynomials with integral coefficients is constructed such that the residue system modulo m of any polynomial with integral coefficients coincides with the residue system of exactly one polynomial of the set* (§§ 2–5). We shall call the polynomials of this set "completely reduced polynomials" modulo m (§ 5). Necessary and sufficient conditions, in terms of the coefficients, that a polynomial be completely reduced, and the number $N(m)$ of such polynomials, are established (§§ 5, 6) by means of the "chain of residual congruences" modulo m (§§ 3, 4), the "signature" $S(m)$ (§ 4), and the "characteristic" $C(m)$ (§ 5). The signature and the characteristic are symbols which depend only on the modulus m and which are readily found for a given m. The (for our purposes) essential properties of the chain of congruences are immediately determined from $S(m)$ and $C(m)$.

* One possible extension of this work, along group theoretic lines, would have points of contact with a long paper by Zsigmondy, Monatshefte für Mathematik und Physik, vol. 7 (1896), pp. 185–289, *Abelsche Gruppen und ihre Anwendung auf die Zahlentheorie;* with Weber, *Lehrbuch der Algebra*, II, 2d edition (1899), pp. 60–68 (Gruppe der Zahlklassen), and others.

The main link between Part I and Part III is established by the italicized passage above, which indicates the existence of a one-to-one correspondence between the completely reduced polynomials modulo m and the complete residue systems modulo m. In Part III, arithmetical sequences are made use of to establish for any given modulus a certain chain of congruences between the elements of a residue system (residual congruences of the second kind) (§§ 10, 11), by means of which necessary and sufficient conditions are derived that a given set of m integers may be the complete residue system, modulo m, of some polynomial with integral coefficients (§ 12). From the "chain of the second kind" are derived a "signature of m of the second kind," $S(m)$ (§ 11) and a "characteristic of m of the second kind," $C(m)$ (§ 12). It is shown that the new signature and characteristic may be denoted by the same symbols as the signature and characteristic introduced in Part I. The new signature and characteristic are therefore again derived directly from the modulus m, and completely determine the nature of the totality of complete residue systems modulo m (§ 13) in a manner entirely parallel to the determination of the totality of completely reduced polynomials in §§ 5, 6. The number of complete residue systems for a given m is again $N(m)$ (§ 13).

There exists a complete isomorphism between Part I and Part III, which permits us to translate a statement concerning either the system of completely reduced polynomials modulo m or the system of complete residue systems modulo m into a statement concerning the other system (§ 14). A résumé (§ 14) lists the main features of this isomorphism, as far as they are used in the present paper.

In checking the literature on the subject, Dickson's History*—particularly vol. 1, ch. 8 (Higher Congruences), and ch. 11 (Greatest Common Divisor)—has been of greatest value. I wish to express a sincere feeling of obligation to Professor Dickson for the assistance which his splendid book has rendered me in this respect.

I. Residual congruences and completely reduced polynomials

§ 1. The number $\mu(m)$

Definition 1: *We denote*† *by* $\mu(m)$, *or, when no ambiguity is possible, by* μ, *the smallest positive integer such that* $[\mu(m)]!$ *is divisible by* m.

We shall need only the following properties of $\mu(m)$:

[*] *History of the Theory of Numbers*, published by the Carnegie Institution of Washington, vol. 1 (1919), vol. 2 (1920).

[†] Compare Neuberg, Mathesis, vol. 27 (1887), p. 68; Nielsen, Nieuw Archief voor Wiskunde, (2), vol. 10 (1913), pp. 100–106; Kempner, American Mathematical Monthly, vol. 25 (1918), p. 204.

$\boldsymbol{m} = \mathbf{1}$; $\mu(1)$ is arbitrarily defined to have the value 0,

$\boldsymbol{m} = \boldsymbol{p}$, p a prime; then $\mu(m) = p$,

$\boldsymbol{m} = \boldsymbol{p}_1 \cdot \boldsymbol{p}_2 \cdots \boldsymbol{p}_\lambda$, $p_1 < p_2 < \cdots < p_\lambda$ primes; then $\mu(m) = p_\lambda$,

$\boldsymbol{m} = \boldsymbol{p}^{\boldsymbol{\gamma}}$, $\boldsymbol{\gamma} < \boldsymbol{p}$; then $\mu(m) = p \cdot \gamma$,

$\boldsymbol{m} = \boldsymbol{p}^{\boldsymbol{\gamma}}$, $\boldsymbol{\gamma} \geqq \boldsymbol{p}$; in this case p^γ is no longer the highest power of p contained in $(p \cdot \gamma)!$ as a factor. Instead, $\mu(m)$ is now determined in the following manner. For any integer $k - 1$ let p^ρ be the highest power of p dividing $(k - 1)!$. The change in the exponent ρ caused by passing from $(k - 1)!$ to $k!$ will be as follows: ρ goes over into $\rho + \tau$ when $k \equiv 0 \bmod p^\tau$, but $k \not\equiv 0 \bmod p^{\tau+1}$; that is, for k not divisible by p, the exponent ρ does not change; for k divisible by p, but not by p^2, ρ increases by unity, etc., Therefore for any prime p and any positive integer k, the highest power of p dividing $k!$ can be found by the following simple scheme:*

Write in a horizontal line, as far as required, the multiples of p: $1 \cdot p$, $2 \cdot p$, $3 \cdot p$, $\cdots$; for every positive integer t, write t under each one of these numbers containing as a factor p^t, but not containing p^{t+1}. For any number $k \cdot p$ in our first row, the exponent of the highest power of p contained as a factor in $(k \cdot p)!$ is obtained by adding all numbers written under $1 \cdot p$, $2 \cdot p$, $\cdots$, $k \cdot p$. This is indicated by the following schedule:

$1\cdot p$	$2\cdot p$	$\cdots$	p^2-p	p^2	p^2+p	$\cdots$	$2p^2-p$	$2p^2$	$2p^2+p$	$\cdots\cdots$	p^3-p	p^3	p^3+p	$\cdots$
1	1	$\cdots$	1	2	1	$\cdots$	1	2	1	$\cdots\cdots$	1	3	1	$\cdots$
1	2	$\cdots$	$p-1$	$p+1$	$p+2$	$\cdots$	$2p$	$2p+2$	$2p+3$	$\cdots\cdots$	p^2+p-2	p^2+p+1	p^2+p+2	$\cdots$.

Example: $p = 3$

1·3	2·3	3·3	4·3	5·3	6·3	7·3	8·3	9·3	10·3	11·3	12·3	13·3	14·3	15·3	$\cdots$
1	1	2	1	1	2	1	1	3	1	1	2	1	1	2	$\cdots$
1	2	4	5	6	8	9	10	13	14	15	17	18	19	21	$\cdots$,

so that, for example, $\mu(3^{10}) = 24$, since $8 \cdot 3$ is the smallest integer such that $24! \equiv 0 \bmod 3^{10}$. Also, $\mu(3^{11}) = \mu(3^{12}) = \mu(3^{13}) = 27$, etc. (For $\gamma < p$, only numbers 1 would occur in the second row.)

$\boldsymbol{m} = \boldsymbol{p}_1^{\gamma_1} \cdot \boldsymbol{p}_2^{\gamma_2} \cdots \boldsymbol{p}_\lambda^{\gamma_\lambda}$, ($m$ any positive integer); in this case $\mu(m)$ is the largest (or one of the largest, in case several of the largest are equal) of the numbers $\mu(p_1^{\gamma_1})$, $\mu(p_2^{\gamma_2})$, $\cdots$, $\mu(p_\lambda^{\gamma_\lambda})$, that is,

$$\mu(m) = \text{Max.}\,(\mu(p_1^{\gamma_1}), \mu(p_2^{\gamma_2}), \cdots, \mu(p_\lambda^{\gamma_\lambda})).$$

From the last case, we derive immediately: *If m_1, m_2 are relatively prime,* $\mu(m_1 \cdot m_2) = \text{Max.}\,(\mu(m_1), \mu(m_2))$.

[*] In Kempner, loc. cit., a formula for $\mu(m)$ is given.

§ 2. Residual congruences modulo m

DEFINITION 2: *We call two polynomials $\phi(x)$, $\psi(x)$, with integral coefficients, residually congruent modulo m, and write*

$$\phi(x) \underline{\equiv} \psi(x) \pmod{m},$$

when $\phi(x) \equiv \psi(x) \pmod{m}$ for all integers x.

*We shall refer to such congruences as residual congruences.** *In particular, $\phi(x) \underline{\equiv} 0 \pmod{m}$ is equivalent to the statement: $\phi(x)$ is a polynomial with integral coefficients which is divisible by m for all integral values of x.*

For convenience, the following facts, most of which are well known, are stated as lemmas.

LEMMA 1: (*a*) *For any positive integer d,*

$$\prod_{i=0}^{\mu(d)-1} (x-i) \underline{\equiv} 0 \pmod{d};$$

(*b*) *For any positive integer d and any multiple m of d*

$$\frac{m}{d} \cdot \prod_{i=0}^{\mu(d)-1} (x-i) \underline{\equiv} 0 \pmod{m}.$$

* An *identical* congruence $\phi(x) \equiv \psi(x) \pmod{m}$, which implies that corresponding coefficients in $\phi(x)$ and $\psi(x)$ are congruent modulo m, is always at the same time a *residual* congruence modulo m; but a *residual* congruence, which only implies that for any integral value of x the residues modulo m of $\phi(x)$ and of $\psi(x)$ have the same value, is not necessarily an *identical* congruence. The types of residual congruences best known are $\Pi_{k=0}^{m-1}(x-k) \underline{\equiv} 0 \pmod{m!}$, and the Fermat congruence $x^p \equiv x \pmod{p}$, p a prime. A partial examination of residual congruences has been made by Borel and Drach, *Introduction à l'étude de la théorie des nombres, etc.*, Paris, 1895, pp. 339–342, who use as a starting point Fermat's theorem, and by Nielsen, loc. cit. (§ 1), whose results are closely related to those of the present § 2, and which are derived by a similar method. Nielsen's object was the determination of the general type of "perfect polynomials," that is, of polynomials $f(x)$ with *rational* coefficients, not all of which are integers, and such that $f(x)$ is an integer for all integral values of x. Nielsen's paper has no connection with work following §§ 1, 2 of the present paper. The arrangement of his work and his notation are not well suited to our purposes, so that a simple reference to the article would not be a satisfactory substitute for this § 2.

A very simple and elegant treatment of the related problem of determining the greatest common divisor d of all integers which can be represented by a *given* polynomial $f(x)$ with integral coefficients and for integral values of x, is due to Hensel, Journal für Mathematik, vol. 116 (1896), pp. 350–356. Hensel proves that d is the greatest common divisor of the $n+1$ numbers $f(0), f(1), \cdots, f(n)$, where n is the degree of the polynomial.

We also mention that, using the methods of Nielsen or of the present paper, it would be more natural to express, for example, necessary and sufficient conditions that a polynomial with integral coefficients be residually congruent to zero for a given modulus, in terms of the coefficients a_i of $f(x) = \Sigma a_i \cdot \binom{x}{i}$, that is, in terms of the coefficients of Newton's interpolation formula, rather than in terms of the coefficients c_i of the polynomial written in the form $f(x) = \Sigma c_i \cdot x^i$. For special purposes, McAtee, American Journal of Mathematics, vol. 41 (1919), pp. 225–242 (239), needed the necessary and sufficient conditions that $c_0 + c_1 x + c_2 x^2 + c_3 x^3 \underline{\equiv} 0 \pmod{4}$, in terms of the c_i, and found them to be: $c_0 \equiv 0$, $2c_1 \equiv 2c_2 \equiv c_1 + c_2 + c_3 \equiv 0 \pmod{4}$. The present paper does not deal with this problem.

Proof: (a) follows from the definition of $\mu(d)$ (see § 1) and from the fact that $\binom{x}{\mu}$ is an integer. (b) is an immediate consequence of (a).

Lemma 2: *Any polynomial* $f(x) = \sum_{k=0}^{k=n} c_k \cdot x^k$ *is uniquely representable in the form** $\sum_{k=0}^{k=n} a_k \cdot \binom{x}{k}$, *and* a_n, c_n *are both different from zero if one of them is different from zero.*

Proof: The actual transformation from one form to the other yields an immediate proof. (See, for example, Hensel, loc. cit. at the end of the first footnote of this section.)

Lemma 3: *If* c_i $(i = 0, 1, \cdots, n)$ *are integers, then* $a_i/i!$ $(i = 0, 1, \cdots, n)$, *and therefore* a_i, *are integers.*

Proof: Compare coefficients of like powers of x in

$$\textstyle\sum_{k=0}^{k=n} c_k \cdot x^k = \sum_{k=0}^{k=n} a_k \cdot \binom{x}{k}.$$

Lemma 4: *If* $f(x) \equiv 0 \ (mod\ m)$, *then* $a_i \equiv 0 \ (mod\ m)$ $(i = 0, 1, \cdots, n)$.

Proof: Let successively $x = 0, 1, \cdots, n$, and consider the resulting recurrence formulæ for a_i.

By Lemma 4, if $f(x) \equiv 0 \pmod m$, and $f(x)$ a polynomial with integral coefficients, $f(x)$ must be of the form

$$\textstyle\sum_{k=0}^{k=n} m\alpha_k \cdot \binom{x}{k} = m\alpha_0 + \sum_{k=1}^{k=n} \left\{ \frac{m\alpha_k}{k!} \cdot \prod_{i=0}^{k-1} (x - i) \right\}, \tag{1}$$

where α_k $(k = 0, 1, \cdots, n)$ are integers (positive, negative, or zero). We consider the coefficients $m\alpha_k/k!$. It follows from the definition of $\mu(m)$ that, considering u, v as unknown positive integers, the smallest value of v, for which $m \cdot u = v!$, is $v = \mu(m)$; and to this value of v corresponds

$$u = \mu(m)!/m.$$

Therefore the smallest value of k for which on the right side of (1) the coefficient $m\alpha_k/k!$ is unity, is $k = \mu(m)$.

We have thus derived:

Lemma 5: *For a given modulus m the residual congruence*

$$1 \cdot \prod_{i=0}^{\mu(m)-1} (x - i) \equiv 0 \pmod m \tag{2}$$

is of lowest possible degree consistent with the condition that the coefficient of the highest power of x shall be unity; or, stating this result in a different form:

For any divisor d of m

$$\frac{m}{d} \cdot \prod_{i=0}^{\mu(d)-1} (x - i) \equiv 0 \pmod m \tag{3}$$

* The fact that this is the expression for the sum of an arithmetical progression of order n is of importance in Part III and (implicitly) in Part I (compare § 10).

is a residual congruence modulo m of lowest possible degree consistent with the condition that the coefficient of the leading term shall have with m the greatest common divisor m/d.

From the preceding we abstract the following

THEOREM I: (*a*) *In any residual congruence modulo m, of degree exactly $\mu(d)$, where d is any divisor of m (including $d = m$ and the trivial case $d = 1$), the coefficient of the leading term*[*] *is either m/d or a multiple of m/d.*

(*b*) *In any residual congruence modulo m, of degree $< \mu(d)$, the coefficient of the leading term*[*] *has with m a greatest common divisor larger than m/d, and is therefore a multiple of m/d.*

For our purposes, the degree and the coefficient of the highest power of x in a residual congruence will be shown to be of particular importance. With this in mind, we introduce the following

DEFINITION 3: *We shall usually indicate a residual cougruence either by*

$$c \cdot x^n \equiv \psi(x) \ (\text{mod } m),$$

where $\psi(x)$ stands for the words[†] *"any polynomial in x of degree $< n$, with integral coefficients, and which is residually congruent to $c \cdot x^n$, modulo m," or, using a still more condensed notation, by $\{n, c\}$, or $\{n, c\}_m$, or $\{n, c\}$ mod m, placing in evidence only the degree n, the coefficient c of the highest power of x, and, where advisable, the modulus m.*

In most cases, the properties just mentioned of $\psi(x) [= \psi_{n-1}(x)]$ are the only ones we shall make use of; consequently, we do not usually think of $\psi(x)$ as being any specific polynomial, but as any polynomial satisfying these conditions.

To illustrate by a simple example: $m = 30$; $1 \cdot x^5 \equiv x$ (mod 30); but also $\Pi_{i=0}^{i=4}(x - i) \equiv 0$ (mod 30); $5x(x-1)(x-2) \equiv 0$ (mod 30); $15 \cdot x(x-1) \equiv 0$ (mod 30). Therefore we may choose in $1 \cdot x^5 \equiv \psi(x)$ (mod 30), $\psi(x) = x$, or $\psi(x) = x + \Pi_{i=0}^{i=4}(x - i)$, or $\psi(x) = x + (ax + b) \cdot 5x(x-1)(x-2) + (dx^2 + ex + f) \cdot 15x(x-1)$, where a, b, d, e, f are any integers; etc. See also Part II.

§ 3. Construction and discussion of the chain of residual congruences modulo m

For a given modulus m we may derive a set of residual congruences by choosing all positive factors of m, including m (and unity, for reasons of convenience) and deriving the corresponding congruence (2), or (3), of § 2. We

[*] "Coefficient of the leading term" = "coefficient of the highest power of the variable which is not $\equiv 0$ (mod m)."

[†] Occasionally we shall indicate the fact that $\psi(x)$ is of degree not higher than $n - 1$, by writing $\psi_{n-1}(x)$.

shall show that these congruences will fall into subsets such that all congruences of any particular subset are implied by a single congruence of the subset. These dependent congruences we shall reject, and the set of congruences retained we shall call a " chain of residual congruences modulo m." We accomplish this by the following process, the discussion of which forms the object of the present and the following paragraph.

CONSTRUCTION OF CHAIN: **1**. Arrange the numbers $\mu(d)$, where d ranges over all positive factors of m (including $m = d_0$ and 1) in order of non-increasing magnitude of $\mu(d)$. Since the $\mu(d)$ are not necessarily all distinct, they will break up into subsets (some or all of which may contain a single element), such that in any subset all $\mu(d)$ are equal.

2. Arrange in each subset the $\mu(d)$ according to decreasing values of d, i.e., according to increasing values of m/d.

3. From each subset, select the first element. Let $\mu(d_i)$ be the element thus selected from the $(i+1)$th subset (where $d_0 = m$), then the numbers $\mu(d_0), \mu(d_1), \cdots, \mu(d_\tau)$, $(d_\tau = 1, \mu(d_\tau) = 0)$, are seen to be arranged so that

$$(a) \quad \mu(d_i) > \mu(d_j), \qquad \text{for} \qquad i < j,$$

$$(b) \quad d_i > d_j, \qquad \frac{m}{d_i} < \frac{m}{d_j}, \qquad \text{for} \qquad i < j.$$

4. For each $\mu(d_i)$ $(i = 0, 1, \cdots, \tau)$ form the residual congruence

$$\frac{m}{d_i} \cdot x^{\mu(d_i)} \equiv \psi(x) \pmod{m}, \qquad \text{i.e.,} \qquad \left\{\mu(d_i), \frac{m}{d_i}\right\}.$$

5. The set of residual congruences modulo m (of which the first is

$$1 \cdot x^{\mu(m)} \equiv \psi(x) \pmod{m},$$

and the last is the trivial congruence $m \cdot x^0 \equiv 0 \pmod{m}$)

$$\frac{m}{d_i} \cdot x^{\mu(d_i)} \equiv \psi(x) \bmod m, \qquad \text{i.e.,} \qquad \left\{\mu(d_i), \frac{m}{d_i}\right\} \quad (i = 0, 1, \cdots, \tau),$$

is called the chain of residual congruences modulo m, or, the chain of congruences modulo m.

This process will be referred to as " Construction, § 3." From the Construction and from Theorem I we read off:

LEMMA 6: *In the chain of congruences as defined above, m/d_i is a proper factor of m/d_{i+1} $(i = 0, 1, \cdots, \tau - 1)$, (and consequently d_{i+1} a proper factor of d_i).*

For the further discussion of the chain we shall need a few facts concerning the types of relation which may exist between two residual congruences for the same modulus.

Let $\{\mu, g\}$ be a residual congruence modulo m, and $(g, m) = c$ the greatest common divisor of m, g. Then the existence of $\{\mu, g\}$ implies the existence of $\{\mu, c\}$, and conversely, as is seen by multiplying both sides of the first congruence by g', where $g \cdot g' \equiv c \pmod{m}$, and both sides of the second congruence by g/c, respectively. For this reason:

We always assume in a residual congruence that the coefficient of the highest power is a factor of the modulus: in $\{\mu, c\} \pmod{m}$, c is always a factor of m (including $c = 1$ and the trivial case $c = m$).

To the end of this paragraph, all residual congruences have the same modulus, m. In a system of two congruences $\{\mu_1, c_1\}$, $\{\mu_2, c_2\}$ we have nine possible combinations of $\mu_1 \gtreqless \mu_2$, $c_1 \gtreqless c_2$. By interchanging indices, these nine cases are reduced to the following five:

1. $c_1 = c_2;\ \mu_1 = \mu_2$. 2. $c_1 = c_2;\ \mu_1 < \mu_2$. 3. $c_1 < c_2;\ \mu_1 = \mu_2$.
4. $c_1 < c_2;\ \mu_1 < \mu_2$. 5. $c_1 < c_2;\ \mu_1 > \mu_2$, or $c_1 > c_2;\ \mu_1 < \mu_2$.

We introduce the following

Definition 4: *Let $I = \{\mu_1, c_1\}$, $II = \{\mu_2, c_2\}$; we call I equivalent to II (or II equivalent to I), and write $I = II$ when both $\mu_1 = \mu_2$, $c_1 = c_2$; and call I stronger than II (or II weaker than I), and write $I > II$ (or $II < I$) when either $c_1 = c_2$, $\mu_1 < \mu_2$; or $c_1 < c_2$, $\mu_1 = \mu_2$; or $c_1 < c_2$, $\mu_1 < \mu_2$. In the cases not accounted for ($c_1 < c_2$, $\mu_1 > \mu_2$; and $c_1 > c_2$, $\mu_1 < \mu_2$) we do not consider I and II as standing to each other in any of the relations*[*] $=, >, <$.

From Definition 4 follows immediately:

Lemma 7: *If $I = II$, I and II are each one implied by the other, and therefore if either one is true, the other is also true.*

If $I > II$, ($II < I$), II is implied by I, and therefore if I, a stronger congruence, is true, II, a weaker congruence, is a fortiori true.

While *any two congruences of the chain for a given m are not in any of the relations $=, >, <$, to each other,* we have

Theorem II: *If we denote by II any residual congruence modulo m then either*

(α) there is a congruence of the chain, I, such that $I = II$; or

(β) there is a congruence of the chain, I, such that $I > II$.

In other words: Given any residual congruence modulo m, the chain contains a congrhence not weaker than it.

[*] We develop the character of the relations between residual congruences only as far as they are needed for this paragraph. The use of the symbols $=, >, <$ in this connection is justified in view of the fact that (1) from $I > II$ follows $II < I$; (2) from $I = II$, $II = III$ follows $I = III$; (3) from $I > II$, $II > III$ follows $I > III$, and similarly for $<$; etc. The results of the remainder of this paragraph may perhaps also be expressed in the language of Kronecker modular systems.

Proof: Let $II = \{\gamma, c\}$ be the congruence to be examined, where c is a factor of m, say $c \cdot \delta = m$. If $\mu(d_0), \mu(d_1), \cdots, \mu(d_\tau)$ are the various values, in descending order of magnitude, which $\mu(d)$ assumes as d ranges over all divisors of m, then we know from " Construction " that our chain contains for each exponent $\mu(d_i)$ a congruence $I = \{\mu(d_i), m/d_i\}$. Therefore, for one of these factors, say for d_i, $\mu(d_i) = \mu(\delta)$. By Construction, either $d_i = \delta$, or $d_i = k \cdot \delta$, $k > 1$ an integer. *Assume first* $\delta = d_i$: By Theorem I, § 2, I is the congruence of lowest possible degree with coefficient m/d_i. Hence, if $\gamma = \mu(d_i)$, $I = II$ by Definition 4; if $\gamma < \mu(d_i)$, II is impossible, by Theorem I; if $\gamma > \mu(d_i)$, $I > II$, by Definition 4. *Assume next* $d_i = k \cdot \delta$, $k > 1$, with $\mu(d_i) = \mu(\delta)$, as above. By Theorem I, $II' = \{\mu(\delta), c\}$ is the congruence of lowest possible degree with coefficient c. By Definition 4, $I > II'$. If $\gamma = \mu(\delta)$, $II = II'$, and hence $I > II$; if $\gamma < \mu(\delta)$, II is impossible; finally, if $\gamma > \mu(\delta)$, $II' > II$, and, since $I > II'$ (above), also $I > II$. Q.e.d.

We have thus established, by combining the Construction with the last theorem:

Theorem II^a: *For every modulus m the Construction of the present section establishes a completely determined chain of residual congruences*

$$\frac{m}{d_i} \cdot x^{\mu(d_i)} \equiv \psi(x) \pmod{m}, \qquad i.e., \qquad \left\{\mu(d_i), \frac{m}{d_i}\right\} \quad (i = 0, 1, \cdots, \tau),$$

and every residual congruence is either equivalent to, or weaker than, a congruence of the chain.

Compare Part II for the relation of the chain to Kronecker modular systems.

§ 4. The signature $S(m)$; continuation of the discussion of the chain

At the present stage of our investigation, we are handicapped by the length of the process described under the Construction of § 3. Our next step consists in showing how the chain may be obtained directly, for a given modulus m, by simple arithmetical operations. This is of importance if the method is to be applicable to a given numerical case.

· From the chain $(m/d_i) \cdot x^{\mu(d_i)} \equiv \psi(x) \pmod{m}$, or $\{\mu(d_i), m/d_i\}$ $(i = 0, 1, \cdots, \tau)$, we derive a symbol which we call the " signature of m," $S(m)$, and which is obtained by placing together, for a given m, all the separate $\{\mu(d_i), m/d_i\}$ $(i = 0, 1, \cdots, \tau)$ in decreasing order of magnitude of d_i, but writing for convenience $\frac{\mu(d_i)}{m/d_i}$ for each i. This leads to

Definition 5: *For a given modulus m, we define the signature S(m) as follows:*

$$S(m) = \begin{bmatrix} \mu(m) & \mu(d_1) & \cdots & \mu(d_{\tau-1}) & \mu(d_\tau) = 0 \\ 1 & \dfrac{m}{d_1} & \cdots & \dfrac{m}{d_{\tau-1}} & \dfrac{m}{1} = m \end{bmatrix},$$

or (ignoring the trivial congruence),

$$S(m) = \begin{bmatrix} \mu(m) & \mu(d_1) & \cdots & \mu(d_{\tau-1}) \\ 1 & \dfrac{m}{d_1} & \cdots & \dfrac{m}{d_{\tau-1}} \end{bmatrix},$$

where $\{\mu(d_i), m/d_i\}$ $(i = 0, 1, \cdots, \tau)$ *is the chain of congruences modulo* m.

It follows that the chain of congruences is completely determined by the signature and conversely, so that our problem reduces to the determination of the signature for a given m. We consider successively the cases: $m = p$, a prime; $m = p_1 \cdot p_2 \cdots p_\tau$, $p_1 < p_2 < \cdots < p_\tau$, primes; $m = p^\gamma$, $\gamma < p$, p a prime; $m = p^\gamma$, $\gamma \geqq p$, p a prime; $m = p_1^{\gamma_1} \cdot p_2^{\gamma_2} \cdots p_\nu^{\gamma_\nu}$, $p_1, \cdots, p_\nu$ distinct primes.

1. $\boldsymbol{m = p}$: the construction of § 3 shows that the chain consists of the single congruence $\{p, 1\}$, besides the trivial $\{0, p\}$. Therefore

$$S(p) = \begin{pmatrix} p & 0 \\ 1 & p \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} p \\ 1 \end{pmatrix}.$$

Confusion with the binomial coefficient is not to be feared.

2. $\boldsymbol{m = p_1 \cdot p_2 \cdots p_\tau}$, $p_1 < p_2 < \cdots < p_\tau$: from § 1, it follows that the factors $d_0, d_1, \cdots, d_\tau$ of step 4 of the Construction, and their respective μ values, are $d_i = p_1 \cdot p_2 \cdots p_{\tau-i}$, $\mu(d_i) = p_{\tau-i}$ $(i = 0, 1, \cdots, \tau)$, so that the residual congruences are $(m/d_i) \cdot x^{\mu(d_i)} \equiv \psi(x) \pmod{m}$, or,

$$\frac{m}{p_1 p_2 \cdots p_{\tau-i}} \cdot x^{p_{\tau-i}} \equiv \psi(x),$$

or $p_{\tau-i+1} \cdot p_{\tau-i+2} \cdots p_\tau \cdot x^{p_{\tau-i}} \equiv \psi(x)$, or $\{p_{\tau-i}, (p_{\tau-i+1} \cdots p_\tau)\}$.

$$S(p_1 p_2 \cdots p_\tau) = \begin{pmatrix} p_\tau & p_{\tau-1} & p_{\tau-2} & \cdots & p_1 & 0 \\ 1 & p_\tau & p_\tau \cdot p_{\tau-1} & \cdots & (p_\tau \cdot p_{\tau-1} \cdots p_2) & (p_\tau \cdot p_{\tau-1} \cdots p_1) \end{pmatrix}$$

$$= \begin{pmatrix} p_\tau & p_{\tau-1} & p_{\tau-2} & \cdots & p_1 \\ 1 & p_\tau & p_\tau \cdot p_{\tau-1} & \cdots & (p_\tau \cdot p_{\tau-1} \cdots p_2) \end{pmatrix}.$$

Example: $m = 2 \cdot 3 \cdot 5 \cdot 11$, $\mu(m) = 11$.
Chain of congruences: $\{11, 1\}$, $\{5, 11\}$, $\{3, 11 \cdot 5\}$, $\{2, 11 \cdot 5 \cdot 3\}$, $\{0, 11 \cdot 5 \cdot 3 \cdot 2\}$;

$$S(2 \cdot 3 \cdot 5 \cdot 11) = \begin{pmatrix} 11 & 5 & 3 & 2 \\ 1 & 11 & 11 \cdot 5 & 11 \cdot 5 \cdot 3 \end{pmatrix}.$$

Written out, the chain may be represented by either one of the systems: (all modulo $2 \cdot 3 \cdot 5 \cdot 11$)

$$\begin{array}{rcr} 1 \cdot x^{11} \equiv \psi(x) & & 1 \cdot \Pi_{i=0}^{i=10} (x - i) \equiv 0 \\ 11 \cdot x^5 \equiv \psi(x) & & 11 \cdot \Pi_{i=0}^{i=4} (x - i) \equiv 0 \\ & \text{or} & \\ 11 \cdot 5 \cdot x^3 \equiv \psi(x) & & 11 \cdot 5 \cdot \Pi_{i=0}^{i=2} (x - i) \equiv 0 \\ 11 \cdot 5 \cdot 3 \cdot x^2 \equiv \psi(x) & & 11 \cdot 5 \cdot 3 \cdot \Pi_{i=0}^{i=1} (x - i) \equiv 0. \end{array}$$

3. $\boldsymbol{m = p^\gamma, \gamma < p}$. Now $d_i = p^{\gamma - i}$ $(i = 0, 1, \cdots, \gamma)$, and the chain of congruences is

$$p^i \cdot x^{p(\gamma - i)} \equiv \psi(x) \pmod{p^\gamma}, \qquad \text{or} \qquad \{p(\gamma - i), p^i\} \quad (i = 0, 1, \cdots, \gamma).$$

Hence

$$S(p^\gamma) = \begin{pmatrix} p\gamma & p(\gamma-1) & p(\gamma-2) & \cdots & p & 0 \\ 1 & p & p^2 & \cdots & p^{\gamma-1} & p^\gamma \end{pmatrix}$$
$$= \begin{pmatrix} p\gamma & p(\gamma-1) & p(\gamma-2) & \cdots & p \\ 1 & p & p^2 & \cdots & p^{\gamma-1} \end{pmatrix}.$$

Example: $m = 5^4$, $\mu(5^4) = 20$,

$$S(5^4) = \begin{pmatrix} 20 & 15 & 10 & 5 \\ 1 & 5 & 5^2 & 5^3 \end{pmatrix}.$$

Chain: $\{20, 1\}$, $\{15, 5\}$, $\{10, 5^2\}$, $\{5, 5^3\}$, i.e.,

$$\begin{array}{lcl} 1 \cdot x^{20} \equiv \psi(x) & & 1 \cdot \Pi_{i=0}^{i=19}(x - i) \equiv 0 \pmod{5^4} \\ 5 \cdot x^{15} \equiv \psi(x) & & 5 \cdot \Pi_{i=0}^{i=14}(x - i) \equiv 0 \pmod{5^4} \\ & \text{or} & \\ 5^2 \cdot x^{10} \equiv \psi(x) & & 5^2 \cdot \Pi_{i=0}^{i=9}(x - i) \equiv 0 \pmod{5^4} \\ 5^3 \cdot x^{5} \equiv \psi(x) & & 5^3 \cdot \Pi_{i=0}^{i=4}(x - i) \equiv 0 \pmod{5^4}. \end{array}$$

4. $\boldsymbol{m = p^\gamma, p \geqq \gamma}$. Now, as in 3, the factors of m are exactly the numbers p^β, $\beta = 0, 1, \cdots, \gamma$; but to different values of β may now correspond the same value of μ (see § 1). From § 1, and from Construction, § 3, it is clear that the values of μ for all divisors of m are exactly the multiples of p up to $\mu(p^\gamma)$, which is itself a multiple of p. In our chain

$$(m/d_i) \cdot x^{\mu(d_i)} \equiv \psi(x), \qquad \text{or} \qquad \left\{\mu(d_i), \frac{m}{d_i}\right\} \quad (i = 0, 1, \cdots, \tau),$$

the degrees*

$$\mu(d_i) = \mu(p^\gamma) - i \cdot p \qquad (i = 0, 1, \cdots, \tau)$$

and the d_i are easily determined by the following schedule (compare § 1)

$$\begin{array}{lccccccc} & \mu(d_{\tau-1}) & \mu(d_{\tau-2}) & \cdots & \mu(d_{\tau-i}) & \cdots & \mu(d_1) & \mu(d_0) \\ & = 1 \cdot p & = 2 \cdot p & & = i \cdot p & & = \mu(p^\gamma) - p & = \mu(p^\gamma) \\ (4) & \delta_1 & \delta_2 & \cdots & \delta_i & \cdots & \delta_{\tau-1} & \delta_\tau \\ & \delta_1 & \delta_1 + \delta_2 & \cdots & \sum_{j=1}^{j=i} \delta_j & \cdots & \sum_{j=1}^{j=\tau-1} \delta_j & \sum_{j=1}^{j=\tau} \delta_j . \end{array}$$

Here, for $i < \tau$, $\delta_i = t + 1$ when $i \equiv 0 \pmod{p^t}$ but $i \not\equiv 0 \pmod{p^{t+1}}$, that is, for $i < \tau$, δ_i is the exponent of the highest power of p dividing

* Where $d_\tau = 1$, $d_{\tau-1} = \mu(p) = p$.

$\mu(d_{\tau-i}) = i \cdot p$ (so that $\delta_1 = \delta_2 = \cdots = \delta_{p-1} = 1$, $\delta_p = 2$, $\delta_{p+1} = \delta_{p+2} = \cdots = \delta_{2p-1} = 1$, $\delta_{2p} = 2$, $\delta_{2p+1} = 1$, $\cdots$, $\delta_{p^2} = 3$, $\delta_{p^2+1} = 1$, $\cdots$. In case 3, all $\delta = 1$; hence the greater complexity of case 4. For $i = \tau$, δ_τ is chosen so that $\sum_{j=i}^{j=\tau} \delta_j = \gamma$. In (4) therefore, for $i < \tau$, $\sum_{j=1}^{j=i} \delta_j$ is the exponent of the highest power of p dividing $[\mu(d_{\tau-i})]!$ so that, by Construction, § 3,

$$p^{\sum_{j=1}^{j=i} \delta_j} = d_{\tau-i} \qquad (i = 1, 2, \cdots, \tau - 1),$$

while

$$p^{\sum_{j=1}^{j=\tau} \delta_j} = p^\gamma = m$$

does not divide $[\mu(d_1)]!$, but does divide $[\mu(d_0)]! = [\mu(m)]!$, without, however, being necessarily the highest power of p dividing $[\mu(m)]!$. (See § 1, Example, and the example below.) Our chain of congruences modulo p^γ, $\{\mu(d_i), m/d_i\}$, therefore is

$$\begin{aligned} 1 \cdot x^{\mu(p^\gamma)} &\equiv \psi(x) \bmod p^\gamma, \\ p^{\delta_\tau} \cdot x^{\mu(p^\gamma)-p} &\equiv \psi(x) \bmod p^\gamma, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ p^{\sum_{j=1}^{j=i} \delta_{\tau-j+1}} \cdot x^{\mu(p^\gamma)-ip} &\equiv \psi(x) \bmod p^\gamma, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ p^{\gamma-1} \cdot x^p &\equiv \psi(x) \bmod p^\gamma, \end{aligned}$$

or*, shorter,

$$\left\{\mu(p^\gamma) - i \cdot p,\ p^{\sum_{j=1}^{j=i} \delta_{\tau-j+1}}\right\} \qquad (i = 0, 1, \cdots, \tau),$$

where δ_i is determined from the schedule (4) above, and where $\sum_{j=1}^{j=i} \delta_{\tau-j+1}$ is defined to be zero for $i = 0$. Hence

$$S(p^\gamma) = \begin{pmatrix} \mu(p^\gamma) & \mu(p^\gamma)-p & \cdots & \mu(p^\gamma)-i\cdot p & \cdots & 2p & p & 0 \\ 1 & p^{\delta_\tau} & \cdots & p^{\delta_\tau+\delta_{\tau-1}+\cdots+\delta_{\tau-i+1}} & \cdots & p^{\delta_\tau+\delta_{\tau-1}+\cdots+\delta_3} & p^{\gamma-1} & p^\gamma \end{pmatrix}.$$

Case 3 is contained in 4 as a special case.

Example: $m = 3^{11}$.

	3	6	9	12	15	18	21	24	27
	1	1	2	1	1	2	1	1	1 †
	1	2	4	5	6	8	9	10	11
$11 - \Sigma\delta_i$:	10	9	7	6	5	3	2	1	0 .

Therefore $\mu(3^{11}) = 27$, and

$$S(3^{11}) = \begin{pmatrix} 27 & 24 & 21 & 18 & 15 & 12 & 9 & 6 & 3 & 0 \\ 1 & 3^1 & 3^2 & 3^3 & 3^5 & 3^6 & 3^7 & 3^9 & 3^{10} & 3^{11} \end{pmatrix},$$

leading to the chain of congruences

* $p^{\delta_\tau+\delta_{\tau-1}+\cdots+\delta_2} = p^{\gamma-1}$, since $\delta_1 = 1$. The trivial congruence, corresponding to $i = \tau$, is here, and frequently elsewhere, suppressed.

† Under 27 we write 1, not 3, because $\Sigma\delta_i$ must be 11. Of course 27! is the smallest factorial of which 3^{11} is a factor, but also 3^{12} and 3^{13} are factors of 27!.

$$1\cdot x^{27}\equiv\psi(x),\ 3\cdot x^{24}\equiv\psi(x),\ \cdots,\ 3^{10}\cdot x^{3}\equiv\psi(x),\ \text{modulo } 3^{11},$$

or,

$$\{27, 1\},\ \{24, 3\},\ \{21, 3^2\},\ \{18, 3^3\},\ \{15, 3^5\},\ \{12, 3^6\},\ \{9, 3^7\},\ \{6, 3^9\},\ \{3, 3^{10}\}.$$

The influence of $\gamma \geqq p$ is reflected in the fact that the exponent in the leading coefficients of the congruences does not range over all values $0, 1, 2, \cdots, 10$, but only over $0, 1, 2, 3, 5, 6, 7, 9, 10$.

We proceed to the general case,

5. $m = p_1^{\gamma_1}\cdot p_2^{\gamma_2}\cdots p_\nu^{\gamma_\nu}$, $p_1, p_2, \cdots, p_\nu$ distinct primes. We derive first a chain for $m = m_1\cdot m_2$, $m_1 = p_1^{\gamma_1}$, $m_2 = p_2^{\gamma_2}$, and therefore $(m_1, m_2) = 1$ In the Construction of § 3, we have to consider $\mu(d)$ for all $d = p_1^{\rho_1}\cdot p_2^{\rho_2}$, $0 \leqq \rho_1 \leqq \gamma_1$, $0 \leqq \rho_2 \leqq \gamma_2$, and then to collect in separate sets all d for which $\mu(d)$ has the same value, and to retain for each μ only the largest d (in case there is one largest; otherwise, one of the largest). From the d so selected we should construct the chain in the following manner: For a $d = p_1^{\rho_1}\cdot p_2^{\rho_2}$ assume $\mu(p_1^{\rho_1}) > \mu(p_2^{\rho_2})$; then $\mu(p_1^{\rho_1}\cdot p_2^{\rho_2}) = \mu(p_1^{\rho_1})$, from § 1, end. Next consider $p_1^{\rho_1}\cdot p_2^{\rho_2+1}$; if $\mu(p_1^{\rho_1}) > \mu(p_2^{\rho_2+1})$, we have $\mu(p_1^{\rho_1}\cdot p_2^{\rho_2+1}) = \mu(p_1^{\rho_1})$, and the factor $p_1^{\rho_1}\cdot p_2^{\rho_2}$ of m is then disregarded. Letting ρ_2 increase in this way from ρ_2, ρ_2+1, $\cdots$ to γ_2, we shall either still have $\mu(p_1^{\rho_1}) > \mu(p_2^{\gamma_2})$, in which case all factors $p_1^{\rho_1}\cdot p_2^{\gamma_2-1}$, $p_1^{\rho_1}\cdot p_2^{\gamma_2-2}$, $\cdots$, $p_1^{\rho_1}$ are disregarded; or, $\mu(p_1^{\rho_1}) \leqq \mu(p_2^{\gamma_2})$, in which case there must be a largest exponent of p_2, (we designate it by $\rho_2+\lambda-1$), for which still $\mu(p_1^{\rho_1}) > \mu(p_2^{\rho_2+\lambda-1})$, but $\mu(p_1^{\rho_1}) \leqq \mu(p_2^{\rho_2+\lambda})$; then $\mu(p_1^{\rho_1}\cdot p_2^{\rho_2+\lambda}) = \mu(p_2^{\rho_2+\lambda})$. We consider in the same way the factors $p_2^{\rho_2+\lambda}\cdot p_1^{\rho_1}$, $p_2^{\rho_2+\lambda}\cdot p_1^{\rho_1+1}$, $p_2^{\rho_2+\lambda}\cdot p_1^{\rho_1+2}$, etc.: assume that $\mu(p_2^{\rho_2+\lambda}) > \mu(p_1^{\rho_1+\sigma-1})$, but $\mu(p_2^{\rho_2+\lambda}) \leqq \mu(p_1^{\rho_1+\sigma})$; then we disregard factors $p_2^{\rho_2+\lambda}\cdot p_1^{\rho_1}$, $\cdots$, $p_2^{\rho_2+\lambda}\cdot p_1^{\rho_1+\sigma-1}$, and retain as next factor $p_2^{\rho_2+\lambda}\cdot p_1^{\rho_1+\sigma}$; and so forth.

In this manner we systematically discover which factors d of m must be retained, and it is clear that they can be arranged in increasing order according to the schedule $\cdots p_1^{\rho_1}\cdot p_2^{\rho_2}, p_1^{\rho_1'}\cdot p_2^{\rho_2'}, \cdots$, where $\rho_1' \geqq \rho_1$, $\rho_2' \geqq \rho_2$, and the sign of equality holds in at most one of the two relations. It is also clear that the factors to be retained are uniquely determined by this process.

Example: $m = 5^4\cdot 7^3$. $\mu(5^4\cdot 7^3) = 3\cdot 7$. In the following rectangular array for all divisors of m the factors which have to be retained for our chain are in heavy type:

	$\mu(7^1) = 7$,	$\mu(7^2) = 14$,	$\mu(7^3) = 21$,
$\boldsymbol{\mu(5^1) = 5}$,	$\boldsymbol{\mu(5^1\cdot 7^1) = 7}$,	$\mu(5^1\cdot 7^2) = 14$,	$\mu(5^1\cdot 7^3) = 21$,
$\mu(5^2) = 10$,	$\boldsymbol{\mu(5^2\cdot 7^1) = 10}$,	$\boldsymbol{\mu(5^2\cdot 7^2) = 14}$,	$\mu(5^2\cdot 7^3) = 21$,
$\mu(5^3) = 15$,	$\mu(5^3\cdot 7^1) = 15$,	$\boldsymbol{\mu(5^3\cdot 7^2) = 15}$,	$\mu(5^3\cdot 7^3) = 21$,
$\mu(5^4) = 20$,	$\mu(5^4\cdot 7^1) = 20$,	$\boldsymbol{\mu(5^4\cdot 7^2) = 20}$,	$\boldsymbol{\mu(5^4\cdot 7^3) = 21}$.

We shall have the chain

$$\begin{aligned}
1\cdot x^{3\cdot 7} &\equiv \psi(x) \bmod 5^4\cdot 7^3,\\
7\cdot x^{4\cdot 5} &\equiv \psi(x) \bmod 5^4\cdot 7^3,\\
5\cdot 7\cdot x^{3\cdot 5} &\equiv \psi(x) \bmod 5^4\cdot 7^3,\\
5^2\cdot 7\cdot x^{2\cdot 7} &\equiv \psi(x) \bmod 5^4\cdot 7^3,\\
5^2\cdot 7^2\cdot x^{2\cdot 5} &\equiv \psi(x) \bmod 5^4\cdot 7^3,\\
5^3\cdot 7^2\cdot x^{1\cdot 7} &\equiv \psi(x) \bmod 5^4\cdot 7^3,\\
5^3\cdot 7^3\cdot x^{1\cdot 5} &\equiv \psi(x) \bmod 5^4\cdot 7^3,
\end{aligned}$$

or in condensed form: $\{21, 1\}$, $\{20, 7\}$, $\{15, 5 \cdot 7\}$, $\{14, 5^2 \cdot 7\}$, $\{10, 5^2 \cdot 7^2\}$, $\{7, 5^3 \cdot 7^2\}$, $\{5, 5^3 \cdot 7^3\}$. The corresponding signature is

$$S(5^4 \cdot 7^3) = \begin{pmatrix} 21 & 20 & 15 & 14 & 10 & 7 & 5 & 0 \\ 1 & 7 & 5 \cdot 7 & 5^2 \cdot 7 & 5^2 \cdot 7^2 & 5^3 \cdot 7^2 & 5^3 \cdot 7^3 & 5^4 \cdot 7^3 \end{pmatrix}.$$

For a modulus m with a large number of factors this method would be very tedious. We show (first without proof) how we may obtain for the example above $S(5^4 \cdot 7^3)$ by inspection from $S(5^4)$ and $S(7^3)$ and shall then prove that this method applies also to the (general) case $m = m_1 \cdot m_2$, $(m_1, m_2) = 1$.

In

$$S(5^4) = \begin{pmatrix} 20 & 15 & 10 & 5 & 0 \\ 5^0 & 5^1 & 5^2 & 5^3 & 5^4 \end{pmatrix}, \qquad S(7^3) = \begin{pmatrix} 21 & 14 & 7 & 0 \\ 7^0 & 7^1 & 7^2 & 7^3 \end{pmatrix},$$

combine the two sets as follows, arranging the exponents in decreasing order of magnitude:

$$\begin{array}{cccccccc} 21 & 20 & 15 & 14 & 10 & 7 & 5 & 0 \\ 1 = 7^0 & 1 = 5^0 & 5^1 & 7^1 & 5^2 & 7^2 & 5^3 & 5^4, 7^3. \end{array}$$

In the second row every 7^k (except 7^0 and 7^3) is spaced between two powers of 5, say 5^γ, 5^δ, $\gamma < \delta$; thus 7^1 between 5^1 and 5^2, 7^2 between 5^2 and 5^3. We multiply 7^k by 5^δ, and replace 7^k by $7^k \cdot 5^\delta$. Similarly, every 5^l (except 5^4) is spaced between two powers of 7, 7^ϵ, 7^ξ, $\epsilon < \xi$. This 5^l is multiplied by 7^ξ, so that 5^l is replaced by $5^l \cdot 7^\xi$. We obtain

$$\begin{array}{cccccccc} 3 \cdot 7 & 4 \cdot 5 & 3 \cdot 5 & 2 \cdot 7 & 2 \cdot 5 & 1 \cdot 7 & 1 \cdot 5 & 0 \\ 1 = 7^0 & 1 = 5^0 & 5^1 & 7^1 & 5^2 & 7^2 & 5^3 & 7^3, 5^4 \\ 1 = 7^0 & 5^0 \cdot 7^1 & 5^1 \cdot 7^1 & 5^2 \cdot 7^1 & 5^2 \cdot 7^2 & 5^3 \cdot 7^2 & 5^3 \cdot 7^3 & 5^4 \cdot 7^3, \end{array}$$

and the first and third lines represent $S(5^4 \cdot 7^3)$.

There is one more point to consider. It may (and frequently will) happen that one of the congruences in the chain for m_1 and one of the congruences in the chain for m_2 will have the same degree.

For example, in $m = 2^7 \cdot 3^4$, $\mu(2^4) = \mu(3^2) = 6$. We have, for this case:

$$S(2^7) = \begin{pmatrix} 8 & 6 & 4 & 2 & 0 \\ 2^0 & 2^3 & 2^4 & 2^6 & 2^7 \end{pmatrix}, \qquad S(3^4) = \begin{pmatrix} 9 & 6 & 3 & 0 \\ 3^0 & 3^2 & 3^3 & 3^4 \end{pmatrix},$$

and, combining (and writing the coefficients in two separate lines),

$$\begin{array}{ccccccc} 9 & 8 & 6 & 4 & 3 & 2 & 0 \\ 3^0 & & 3^2 & & 3^3 & & 3^4 \\ & 2^0 & 2^3 & 2^4 & & 2^6 & 2^7 \\ \hline 3^0 \cdot 2^0 & 2^0 \cdot 3^2 & 3^2 \cdot 2^3 & 2^4 \cdot 3^3 & 3^3 \cdot 2^6 & 2^6 \cdot 3^4 & 3^4 \cdot 2^7; \end{array}$$

hence

$$S(2^7 \cdot 3^4) = \begin{pmatrix} 9 & 8 & 6 & 4 & 3 & 2 & 0 \\ 1 & 3^2 & 2^3 \cdot 3^2 & 2^4 \cdot 3^3 & 2^6 \cdot 3^3 & 2^6 \cdot 3^4 & 2^7 \cdot 3^4 \end{pmatrix},$$

leading to the chain $\{9, 1\}$, $\{8, 3^2\}$, $\{6, 2^3 \cdot 3^2\}$, $\cdots$, $\{2, 2^6 \cdot 3^4\}$.

We show that the method used in these two examples applies without change to the general case $m = n \cdot n'$, n, n' any factors of m, which we assume to be relatively prime, as we may do without causing loss of generality, since the case $m = p^\gamma$ has already been treated.

For this, we prove

LEMMA 8: *Let*

$$S(n) = \begin{bmatrix} \mu(n) & \mu(d_1) & \cdots & \mu(d_\lambda) \\ 1 = \dfrac{n}{n} & \dfrac{n}{d_1} & \cdots & \dfrac{n}{d_\lambda} \end{bmatrix},$$

$$S(n') = \begin{bmatrix} \mu(n') & \mu(d'_1) & \cdots & \mu(d'_\tau) \\ 1 = \dfrac{n'}{n'} & \dfrac{n'}{d'_1} & \cdots & \dfrac{n'}{d'_\tau} \end{bmatrix},$$

so that the d_i ($d_i > d_{i+1}$) are certain factors of n, and the d'_j ($d'_j > d'_{j+1}$), certain factors of n'; assume also that, after all exponents in both signatures have been arranged in decreasing order of magnitude, a part of the combination is as follows:

$$\begin{matrix} \cdots & \mu(d_\alpha) & \mu(d_{\alpha+1}) & \cdots & \mu(d_{\alpha'}) & \mu(d'_\beta) & \mu(d'_{\beta+1}) & \cdots & \mu(d'_{\beta'}) & \mu(d_\gamma) & \mu(d_{\gamma+1}) & \cdots \\ \cdots & \dfrac{n}{d_\alpha} & \dfrac{n}{d_{\alpha+1}} & \cdots & \dfrac{n}{d_{\alpha'}} & \dfrac{n'}{d'_\beta} & \dfrac{n'}{d'_{\beta+1}} & \cdots & \dfrac{n'}{d'_{\beta'}} & \dfrac{n}{d_\gamma} & \dfrac{n}{d_{\gamma+1}} & \cdots, \end{matrix}$$

then the corresponding part of $S(n \cdot n')$ is:

$$\left[\begin{matrix} \cdots & \mu(d_\alpha \cdot d'_\beta) & \mu(d_{\alpha+1} \cdot d'_\beta) & \cdots & \mu(d_{\alpha'} \cdot d'_\beta) & \mu(d'_\beta \cdot d_\gamma) & \mu(d'_{\beta+1} \cdot d_\gamma) & \cdots \\ \cdots & \dfrac{n}{d_\alpha} \cdot \dfrac{n'}{d'_\beta} & \dfrac{n}{d_{\alpha+1}} \cdot \dfrac{n'}{d'_\beta} & \cdots & \dfrac{n}{d_{\alpha'}} \cdot \dfrac{n'}{d'_\beta} & \dfrac{n'}{d'_\beta} \cdot \dfrac{n}{d_\gamma} & \dfrac{n'}{d'_{\beta+1}} \cdot \dfrac{n}{d_\gamma} & \cdots \end{matrix}\right.$$

$$\left.\begin{matrix} \mu(d'_{\beta'} \cdot d_\gamma) & \mu(d_\gamma \cdot d'_\delta) & \mu(d_{\gamma+1} \cdot d'_\delta) & \cdots \\ \dfrac{n'}{d'_{\beta'}} \cdot \dfrac{n}{d_\gamma} & \dfrac{n}{d_\gamma} \cdot \dfrac{n'}{d'_\delta} & \dfrac{n}{d_{\gamma+1}} \cdot \dfrac{n'}{d'_\delta} & \cdots \end{matrix}\right].$$

Proof: For any $\mu(d_i)$ the signature $S(n)$ yields a residual congruence $\{\mu(d_i), n/d_i\}_n$, and for any $\mu(d'_j)$ we get from $S(n')$ a congruence $\{\mu(d'_j), n'/d'_j\}_{n'}$. After combining the signatures in decreasing order of the μ, as assumed in the lemma, we select any one of the $\mu(d)$, for example $\mu(d_i)$. This $\mu(d_i)$ will be followed* by some $\mu(d')$, say $\mu(d'_j)$, and therefore $\mu(d_i) > \mu(d'_j)$, and, since d_i and d'_j are relatively prime, $\mu(d_i) = \mu(d_i d'_j)$ (see § 1, end), where $d_i d'_j$ is some divisor of $n \cdot n'$. Therefore (§ 2), there exists a congruence $\{\mu(d_i d'_j), nn'/d_i d'_j\}_{nn'}$. The possibility of a $\mu(d)$ equaling a $\mu(d')$ offers no difficulty. (See Example $S(2^7 \cdot 3^4)$ above.) It is seen that all residual congruences modulo nn' which the construction of § 3 calls for, actually are obtained in this way. The lemma is thus proved.

* Unless $\mu(d_i)$ is smaller than any $\mu(d')$, which will always happen for the smallest $\mu(d)$, except when the smallest $\mu(d')$ is smaller than any $\mu(d)$. This does not give rise to any difficulty. See Example $S(5^4 \cdot 7^3)$, above.

Since, by assumption, $\mu(d_\alpha) > \mu(d'_\beta)$, $\mu(d_{\alpha+1}) > \mu(d'_\beta)$, $\cdots$, $\mu(d_{\alpha'}) > \mu(d'_\beta)$, we shall have $\mu(d_\alpha \cdot d'_\beta) = \mu(d_\alpha)$, $\mu(d_{\alpha+1} \cdot d'_\beta) = \mu(d_{\alpha+1})$, $\cdots$, $\mu(d_{\alpha'} \cdot d'_\beta) = \mu(d_{\alpha'})$; similarly, $\mu(d'_\beta \cdot d_\gamma) = \mu(d'_\beta)$, $\mu(d'_{\beta+1} \cdot d_\gamma) = \mu(d'_{\beta+1})$, $\cdots$, $\mu(d'_{\beta'} \cdot d_\gamma) = \mu(d'_{\beta'})$; and $\mu(d_\gamma \cdot d'_\delta) = \mu(d_\gamma)$, $\mu(d_{\gamma+1} \cdot d'_\delta) = \mu(d_{\gamma+1})$, $\cdots$. Then part of our signature $S(n \cdot n')$ is given by

$$\begin{matrix} \cdots & \mu(d_\alpha) & \mu(d_{\alpha+1}) & \cdots & \mu(d_{\alpha'}) & \mu(d'_\beta) & \mu(d'_{\beta+1}) & \cdots & \mu(d'_{\beta'}) & \mu(d_\gamma) & \mu(d_{\gamma+1}) & \cdots \\ \cdots & \dfrac{nn'}{d_\alpha\, d'_\beta} & \dfrac{nn'}{d_{\alpha+1}\, d'_\beta} & \cdots & \dfrac{nn'}{d_{\alpha'}\, d'_\beta} & \dfrac{nn'}{d'_\beta\, d_\gamma} & \dfrac{nn'}{d'_{\beta+1}\, d_\gamma} & \cdots & \dfrac{nn'}{d'_{\beta'}\, d_\gamma} & \dfrac{nn'}{d_\gamma\, d'_\delta} & \dfrac{nn'}{d_{\gamma+1}\, d'_\delta} & \cdots \end{matrix}.$$

But this represents exactly the law indicated in the special examples $m = 5^4 \cdot 7^3$, $m = 2^7 \cdot 3^4$ above. To make this perfectly clear, we rewrite a few lines from these problems, with only obvious modifications:

$$n = 5^4, \qquad n' = 7^3. \qquad S(5^4) = \begin{pmatrix} \substack{\mu(5^4)\\=20} & \substack{\mu(5^3)\\=15} & \substack{\mu(5^2)\\=10} & \substack{\mu(5^1)\\=5} & \substack{\mu(1)\\=0} \\ 5^0 & 5^1 & 5^2 & 5^3 & 5^4 \end{pmatrix};$$

$$S(7^3) = \begin{pmatrix} \substack{\mu(7^3)\\=21} & \substack{\mu(7^2)\\=14} & \substack{\mu(7^1)\\=7} & \substack{\mu(1)\\=0} \\ 7^0 & 7^1 & 7^2 & 7^3 \end{pmatrix}.$$

From these, by Lemma 8 and taking the last remarks into account,

$$S(5^4 \cdot 7^3) = \begin{pmatrix} \substack{\mu(7^3\cdot 5^4)\\=\mu(7^3)} & \substack{\mu(5^4\cdot 7^2)\\=\mu(5^4)} & \substack{\mu(5^3\cdot 7^2)\\=\mu(5^3)} & \substack{\mu(7^2\cdot 5^2)\\=\mu(7^2)} & \substack{\mu(5^2\cdot 7^1)\\=\mu(5^2)} & \substack{\mu(7^1\cdot 5^1)\\=\mu(7^1)} & \substack{\mu(5^1\cdot 7^0)\\=\mu(5^1)} \\ 7^0 \cdot 5^0 & 5^0 \cdot 7^1 & 5^1 \cdot 7^1 & 7^1 \cdot 5^2 & 5^2 \cdot 7^2 & 7^2 \cdot 5^3 & 5^3 \cdot 7^3 \end{pmatrix},$$

in agreement with the example.

To show that, here again, the method requires no real modification when a $\mu(d_i)$ equals a $\mu(d_j)$, we indicate in the same manner the work for the other example, $m = 2^7 \cdot 3^4$.

$$n = 2^7, \qquad n' = 3^4. \qquad S(2^7) = \begin{pmatrix} \substack{\mu(2^7)\\=8} & \substack{\mu(2^4)\\=6} & \substack{\mu(2^3)\\=4} & \substack{\mu(2^1)\\=2} & \substack{\mu(2^0)\\=0} \\ 2^0 & 2^3 & 2^4 & 2^6 & 2^7 \end{pmatrix},$$

$$\mu(3^4) = \begin{pmatrix} \substack{\mu(3^4)\\=9} & \substack{\mu(3^2)\\=6} & \substack{\mu(3^1)\\=3} & \substack{\mu(3^0)\\=0} \\ 3^0 & 3^2 & 3^3 & 3^4 \end{pmatrix}.$$

Then

$$S(2^7 \cdot 3^4) = \begin{pmatrix} \substack{\mu(3^4\cdot 2^7)\\=\mu(3^4)} & \substack{\mu(2^7\cdot 3^2)\\=\mu(2^7)} & \substack{\mu(2^4\cdot 3^2)\\=\mu(2^4)=\mu(3^2)} & \substack{\mu(3^2\cdot 2^3)\\=\mu(3^2)} & \substack{\mu(2^3\cdot 3^1)\\=\mu(2^3)} & \substack{\mu(3^1\cdot 2^1)\\=\mu(3^1)} & \substack{\mu(2^1\cdot 3^0)\\=\mu(2^1)} \\ 3^0 \cdot 2^0 & 2^0 \cdot 3^2 & 3^2 \cdot 2^3 & 3^2 \cdot 2^4 & 2^4 \cdot 3^3 & 3^3 \cdot 2^6 & 2^6 \cdot 3^4 \end{pmatrix},$$

i.e.,

$$\begin{pmatrix} 9 & 8 & 6 & 6 & 4 & 3 & 2 \\ 1 & 3^2 & 2^3 \cdot 3^2 & 3^2 \cdot 2^4 & 2^4 \cdot 3^3 & 3^3 \cdot 2^6 & 2^6 \cdot 3^4 \end{pmatrix},$$

in agreement with the example, since $\{6, 3^2 \cdot 2^4\}_{2^7 \cdot 3^4} < \{6, 3^2 \cdot 2^3\}_{2^7 \cdot 3^4}$, and may therefore be omitted.

The general case, $m = p_1^{\gamma_1} \cdot p_2^{\gamma_2} \cdots p_\nu^{\gamma_\nu}$, can always be treated by successive applications of the lemma last proved. However, it would be a very simple matter to display a schematic arrangement by means of which $S(p_1^{\gamma_1} \cdot p_2^{\gamma_2} \cdots p_\nu^{\gamma_\nu})$ is obtained in one step from $S(p_i^{\gamma_i})$ $(i = 1, 2, \cdots, \nu)$, if this were worth while. One example, below, is worked out in this way, and this will suffice to make clear the general process.

Example: $m = 2^7 \cdot 3^4 \cdot 7^2$. $\mu(7^2) = 14$, $\mu(3^4) = 9$, $\mu(2^7) = 8$; $\mu(2^7 \cdot 3^4 \cdot 5^3) = 14$.

$$S(2^7) = \begin{pmatrix} 8 & 6 & 4 & 2 & 0 \\ 2^0 & 2^3 & 2^4 & 2^6 & 2^7 \end{pmatrix}; \qquad S(3^4) = \begin{pmatrix} 9 & 6 & 3 & 0 \\ 3^0 & 3^2 & 3^3 & 3^4 \end{pmatrix} \qquad \text{(see above).}$$

$$S(7^2) = \begin{pmatrix} 14 & 7 & 0 \\ 7^0 & 7^1 & 7^2 \end{pmatrix}.$$

$$\begin{array}{ccccccccc}
14 & 9 & 8 & 7 & 6 & 4 & 3 & 2 & 0 \\
7^0 & & & 7^1 & & & & & 7^2 \\
& 3^0 & & & 3^2 & & 3^3 & & 3^4 \\
& & 2^0 & & 2^3 & 2^4 & & 2^6 & 2^7 \\
\hline
7^0; & 3^0 \cdot 7^1; & 7^1 \cdot 3^2; & 7^1\, 3^2\, 2^3; & 2^3\, 3^2\, 7^2; & 2^4\, 3^3\, 7^2; & 3^3\, 2^6\, 7^2; & 2^6\, 3^4\, 7^2; & 2^7\, 3^4\, 7^2;
\end{array}$$

$S(2^7 \cdot 3^4 \cdot 7^2)$

$$= \begin{pmatrix} 14 & 9 & 8 & 7 & 6 & 4 & 3 & 2 & 0 \\ 1 & 7 & 3^2 \cdot 7 & 2^3 \cdot 3^2 \cdot 7 & 2^3 \cdot 3^2 \cdot 7^2 & 2^4 \cdot 3^3 \cdot 7^2 & 2^6 \cdot 3^3 \cdot 7^2 & 2^5 \cdot 3^4 \cdot 7^2 & 2^7 \cdot 3^4 \cdot 7^2 \end{pmatrix}.$$

Chain: $\{14, 1\}$, $\{9, 7\}$, $\{8, 3^2 \cdot 7\}$, $\{7, 2^3 \cdot 3^2 \cdot 7\}$, $\cdots$, $\{2, 2^5 \cdot 3^4 \cdot 7^2\}$.

We combine the results of §§3, 4 in the following theorem.

THEOREM III: *To any positive integer m (the modulus)—which we assume given in factored form—we can determine by a definite and very simple arithmetical process the signature*

$$S(m) = \begin{bmatrix} \mu(m) & \mu(d_1) & \cdots & \mu(d_i) & \mu(d_{i+1}) & \cdots & \mu(d_{\tau-1}) & \mu(d_\tau) = 0 \\ \dfrac{m}{d_0} = 1 & \dfrac{m}{d_1} & \cdots & \dfrac{m}{d_i} & \dfrac{m}{d_{i+1}} & \cdots & \dfrac{m}{d_{\tau-1}} & \dfrac{m}{d_\tau} = m \end{bmatrix},$$

where each d_i is a proper factor of all d_j, $j > i$, and consequently a divisor of m, and $d_0 = m$, $d_\tau = 1$.

$S(m)$ completely determines the chain of residual congruences modulo m,

$$\frac{m}{d_i} \cdot \prod_{i=0}^{\mu(d_i)-1} (x - i) \equiv 0 \pmod{m}$$

$$(i = 0, 1, \cdots, \tau - 1) \text{ or } (i = 0, 1, \cdots, \tau)$$

or,

$$\frac{m}{d_i} \cdot x^{\mu(d_i)} \equiv \psi(x) \pmod{m} \qquad (i = 0, 1, \cdots, \tau),$$

or, as we also write,[*]

$$\left\{ \mu(d_i), \frac{m}{d_i} \right\} \qquad (i = 0, 1, \cdots, \tau).$$

Every residual congruence modulo m is either equivalent to, or weaker than, a congruence of the chain.

If $m = p_1^{\gamma_1} \cdot p_2^{\gamma_2} \cdots p_\nu^{\gamma_\nu}$, then the number of congruences in the chain modulo m is $\leqq \gamma_1 + \gamma_2 + \cdots + \gamma_\gamma$ (besides the trivial congruence).

[*] When the modulus m is to be explicitly stated,

$$\left\{ \mu(d_i), \frac{m}{d_i} \right\}_m \quad \text{or} \quad \left\{ \mu(d_i), \frac{m}{d_i} \right\} \pmod{m}.$$

§ 5. Completely reduced polynomials modulo m; the characteristic $C(m)$

We make use of the chain of congruences of §§ 3, 4 to reduce any polynomial with integral coefficients to a normal form which is to be, modulo m, residually congruent to the given polynomial. First we may, by repeated application of the first congruence $\{\mu(m), 1\}$, reduce a polynomial of any degree to one of degree at most $\mu(m) - 1$; then repeated application of the second congruence, $\{\mu(d_1), m/d_1\}$, permits us to reduce in the new polynomial the coefficients of $x^{\mu(m)-1}, x^{\mu(m)-2}, \cdots, x^{\mu(d_1)}$ to numbers of the set $0, 1, \cdots, (m/d_1) - 1$. In the same way, $\{\mu(d_2), m/d_2\}$ reduces, without affecting coefficients already reduced, the coefficients of $x^{\mu(d_1)-1}, \cdots, x^{\mu(d_2)}$ to numbers of the set $0, 1, \cdots, (m/d_2) - 1$. Continuing in this manner, we find that finally the coefficients of $x^{\mu(d_{\tau-2})-1}, x^{\mu(d_{\tau-2})-2}, \cdots, x^{\mu(d_{\tau-1})}$ are restricted to the numbers $0, 1, \cdots, (m/d_{\tau-1}) - 1$, while the coefficients of $x^{\mu(d_{\tau-1})-1}, x^{\mu(d_{\tau-1})-2}, \cdots, x^1, x^0$ are not restricted at all beyond the obvious limitation that they shall belong to the set $0, 1, \cdots, m - 1$.

We are thus led to an arrangement of the following type: any polynomial is reducible by the chain of residual congruences modulo m to a normal form in which the degree is $\mu(m) - 1$* and in which

the terms with exponents:	$\mu(1)=0, \cdots, \mu(d_{\tau-1})-1$	$\mu(d_{\tau-1}) \cdots \mu(d_{\tau-2}) - 1$	$\cdots\cdots$	$\mu(d_2) \cdots \mu(d_1) - 1$	$\mu(d_1) \cdots \mu(m) - 1$
have their coefficients restricted to the numbers	$0 \cdots m-1$	$0 \cdots (m/d_{\tau-1}) - 1$	$\cdots\cdots$	$0 \cdots (m/d_2) - 1$	$0 \cdots (m/d_1) - 1.$

We have then: counting the degree of the reduced polynomial as exactly $\mu(m) - 1$, and assuming it written in the form $a_0 + a_1 x + a_2 x^2 + \cdots + a_{\mu(m)-1} \cdot x^{\mu(m)-1}$,

the $\mu(d_{\tau-1})$ lowest coefficients† $a_0, \cdots, a_{\mu(d_{\tau-1})-1}$ are each one of the m numbers $0, 1, \cdots, m-1$;

. . . .

the $\mu(d_{i-1}) - \mu(d_i)$ coefficients $a_{\mu(d_i)}, \cdots, a_{\mu(d_{i-1})-1}$ are each one of the m/d_i numbers $0, 1, \cdots, (m/d_i) - 1$;

. . . .

the $\mu(d_1) - \mu(d_2)$ coefficients $a_{\mu(d_2)}, \cdots, a_{\mu(d_1)-1}$ are each one of the m/d_2 numbers $0, 1, \cdots, (m/d_2) - 1$; and, finally,

the $\mu(m) - \mu(d_1)$ highest coefficients $a_{\mu(d_1)}, \cdots, a_{\mu(m)-1}$ are each one of the m/d_1 numbers $0, 1, \cdots, (m/d_1) - 1$.

* Admitting that one, or more, of the highest powers of x may have a coefficient 0; or all coefficients may be zero, including the constant term.

† For conformity with what follows, we should write: "the $\mu(d_{\tau-1}) - \mu(d_\tau)$ lowest coefficients," but $d_\tau = 1$, $\mu(d_\tau) = 0$ (§ 1, beginning).

This is indicated by the following symbol which we call the characteristic of m, and denote by $C(m)$.

DEFINITION 6: *We define the characteristic of m, $C(m)$, by writing in a first line the number of coefficients in each subset and in a second line the number of values over which the coefficients in each subset may range, and arranging so that going from left to right corresponds to increasing degrees of the terms:*

$$C(m) = \left(\begin{array}{c|c|c|c} \mu(d_{\tau-1}) & \mu(d_{\tau-2}) - \mu(d_{\tau-1}) & \cdots & \mu(d_{i-1}) - \mu(d_i) \\ m & \dfrac{m}{d_{\tau-1}} & \cdots & \dfrac{m}{d_i} \end{array}\right.$$
$$\left.\begin{array}{c|c|c} \cdots & \mu(d_1) - \mu(d_2) & \mu(m) - \mu(d_1) \\ \cdots & \dfrac{m}{d_2} & \dfrac{m}{d_1} \end{array}\right).$$

We call a polynomial of degree $\leqq \mu(m) - 1$, and of which the coefficients are restricted as indicated, a completely reduced polynomial modulo m, or simply a completely reduced polynomial.

Two completely reduced polynomials modulo m cannot be residually congruent modulo m without being identical; that is, if $f_1(x) \equiv f_2(x) \bmod m$, and $f_1(x)$, $f_2(x)$ are both completely reduced modulo m, then $f_1(x)$ and $f_2(x)$ have corresponding coefficients equal to each other. For, if this were not so, we should have $\phi(x) = f_1(x) - f_2(x) \equiv 0 \bmod m$; and in the polynomial $\phi(x)$ at least one coefficient would be different from zero. Assume that in $\phi(x)$ the highest term whose coefficient does not vanish is $c \cdot x^\gamma$, $\mu(d_i) \leqq \gamma < \mu(d_{i-1})$. Since both in $f_1(x)$ and in $f_2(x)$ the coefficient of x^γ lies between 0 (incl.) and m/d_i (excl.), we shall have $0 \leqq c < m/d_i$, thus contradicting Theorem 1, § 2. In particular, a completely reduced polynomial is residually congruent to zero when and only when all coefficients and the constant term are zero.

We have thus proved

THEOREM IV: *Every polynomial with integral coefficients is modulo m residually congruent to one and only one completely reduced polynomial. In greater detail:*

If the signature of m is

$$S(m) = \begin{pmatrix} \mu(m) & \mu(d_1) & \mu(d_2) & \cdots & \mu(d_{\tau-1}) & \mu(d_\tau) = 0 \\ 1 & m/d_1 & m/d_2 & \cdots & m/d_{\tau-1} & m/d_\tau = m \end{pmatrix},$$

then the characteristic is

$$C(m) = \left(\begin{array}{c|c|c|c|c} \mu(d_{\tau-1}) & \mu(d_{\tau-2}) - \mu(d_{\tau-1}) & \cdots & \mu(d_1) - \mu(d_2) & \mu(m) - \mu(d_1) \\ m & m/d_{\tau-1} & \cdots & m/d_2 & m/d_1 \end{array}\right),$$

and every polynomial with integral coefficients is modulo m residually congruent to exactly one polynomial $a_0 + a_1 x + \cdots + a_{\mu(d_i)-1} \cdot x^{\mu(d_i)-1} + a_{\mu(d_i)} \cdot x^{\mu(d_i)}$

$+ \cdots + a_{\mu(d_{i-1})-1} \cdot x^{\mu(d_{i-1})-1} + a_{\mu(d_{i-1})} \cdot x^{\mu(d_{i-1})} + \cdots + a_{\mu(d_1)-1} \cdot x^{\mu(d_1)-1} + a_{\mu(d_1)} \cdot x^{\mu(d_1)} + \cdots + a_{\mu(m)-1} \cdot x^{\mu(m)-1}$, *in which each of the coefficients* $a_0, \cdots, a_{\mu(d_{\tau-1})-1}$ *has a value* $0, 1, \cdots, m-1$; *each of the coefficients* $a_{\mu(d_i)}, \cdots, a_{\mu(d_{i-1})-1}$ *has a value* $0, 1, \cdots, (m/d_i) - 1$.

For any m, *the characteristic may be immediately read off from the signature.*

When m contains only distinct prime factors, or when m is of the form p^γ, $\gamma < p$, the characteristic is expressible in terms of the prime factors and the exponent, without making explicit use of the μ function:

$$(a)\quad m = p; \qquad S(p) = \begin{pmatrix} p & 0 \\ 1 & p \end{pmatrix}; \qquad C(p) = \left(\begin{array}{|c|} p \\ p \end{array}\right).$$

$$(b)\quad m = p_1 \cdot p_2 \cdots p_\rho, \qquad p_1 < p_2 < \cdots < p_\rho;$$

$$S(p_1 p_2 \cdots p_\rho)$$

$$= \begin{pmatrix} p_\rho & p_{\rho-1} & p_{\rho-2} & \cdots & .p_2 & p_1 & 0 \\ 1 & p_\rho & p_\rho \cdot p_{\rho-1} & \cdots & p_\rho \cdot p_{\rho-1} \cdots p_3 & p_\rho \cdot p_{\rho-1} \cdots p_2 & p_\rho \cdot p_{\rho-1} \cdots p_1 \end{pmatrix};$$

$$C(p_1 p_2 \cdots p_\rho)$$

$$= \left(\begin{array}{c|c|c|c|c|c} p_1 & p_2 - p_1 & p_3 - p_2 & \cdots & p_{\rho-1} - p_{\rho-2} & p_\rho - p_{\rho-1} \\ p_\rho p_{\rho-1} \cdots p_1 & p_\rho \cdot p_{\rho-1} \cdots p_2 & p_\rho p_{\rho-1} \cdots p_3 & \cdots & p_\rho p_{\rho-1} & p_\rho \end{array}\right).$$

$$(c_1)\quad m = p^\gamma, \qquad \boldsymbol{\gamma < p}.$$

$$S(p^\gamma) = \begin{pmatrix} p\gamma & p(\gamma-1) & p(\gamma-2) & \cdots & p^2 & p & 0 \\ 1 & p & p^2 & \cdots & p^{\gamma-2} & p^{\gamma-1} & p^\gamma \end{pmatrix};$$

$$C(p^\gamma) = \left(\begin{array}{c|c|c|c|c|c} p & p & p & \cdots & p & p \\ p^\gamma & p^{\gamma-1} & p^{\gamma-2} & \cdots & p^2 & p \end{array}\right).$$

But whatever the type of m, the characteristic, and thereby the set of completely reduced polynomials modulo m, are in all cases easily determined. Examples are given in § 6.

We see that the chain of residual congruences modulo m, and likewise, the characteristic $C(m)$ or the signature $S(m)$, determine:

1. The structure of the system of completely reduced polynomials modulo m;
2. The structure of the individual completely reduced polynomial modulo m.

§ 6. The number $N(m)$ of completely reduced polynomials modulo m. Classes of polynomials

From the characteristic $C(m)$ it is easily possible to determine the number $N(m)$ of distinct completely reduced polynomials modulo m, since the characteristic gives the degree $\mu(m) - 1$ of the polynomial as well as the range of values for each coefficient. We obtain,* since there are $\mu(m) - \mu(d_1)$ coef-

* Counting as a completely reduced polynomial also the polynomial all of whose coefficients are zero, as well as the constants $1, 2, \cdots, m-1$.

ficients which may independently assume any of the values $0, 1, \cdots, (m/d_1) - 1$, etc.,

THEOREM* V:

$$N(m) = \left(\frac{m}{d_1}\right)^{\mu(m)-\mu(d_1)} \cdot \left(\frac{m}{d_2}\right)^{\mu(d_1)-\mu(d_2)} \cdots \left(\frac{m}{d_{\tau-1}}\right)^{\mu(d_{\tau-2})-\mu(d_{\tau-1})} \cdot m^{\mu(d_{\tau-1})}$$

$$(5) \qquad = m^{\mu(m)} \cdot d_1^{\mu(d_1)-\mu(m)} \cdot d_2^{\mu(d_2)-\mu(d_1)} \cdots d_{\tau-1}^{\mu(d_{\tau-1})-\mu(d_{\tau-2})}$$

$$= \left(\frac{m}{d_1}\right)^{\mu(m)} \cdot \left(\frac{d_1}{d_2}\right)^{\mu(d_1)} \cdots \left(\frac{d_{\tau-2}}{d_{\tau-1}}\right)^{\mu(d_{\tau-2})} \cdot \left(\frac{d_{\tau-1}}{1}\right)^{\mu(d_{\tau-1})}.$$

Special cases are:

(*a*) $m = p; \qquad N(p) = p^p.$

(*b*) $m = p_1 \cdot p_2 \cdots p_\gamma, p_1, \cdots, p_\gamma$ distinct primes:

$$N(p_1 \cdot p_2 \cdots p_\gamma) = p_1^{p_1} \cdot p_2^{p_2} \cdots p_\gamma^{p_\gamma}.$$

(c_1) $m = p^\gamma$, $\gamma < p$; $N(p^\gamma) = p^{p(1+2+\cdots+\gamma)} = p^{p\gamma(\gamma+1)/2}$.

From the manner in which in § 4 the signature of $m_1 \cdot m_2$ was derived from $S(m_1)$ and $S(m_2)$, it follows that, for m_1, m_2 relatively prime, we shall have $N(m_1 \cdot m_2) = N(m_1) \cdot N(m_2)$. However, since we have explained a direct and simple method for determining $N(m)$ for any given m, we shall not discuss the functional properties of $N(m)$.

Our work suggests a division of all polynomials with integral coefficients into $N(m)$ *classes modulo m by grouping always into one class the infinitude of polynomials which are residually congruent modulo m to the same completely reduced polynomial.* All polynomials belonging to the same class will be residually congruent to each other, modulo m, and we may select the completely reduced polynomial as representative of the entire class. The main property of the classes, for our present purposes, is expressed in the (now obvious)

THEOREM VI: *There are exactly $N(m)$ classes of polynomials modulo m; every polynomial belongs to exactly one class; all polynomials belonging to the same class have the same complete residue system modulo m; no two polynomials belonging to distinct classes have the same complete residue system.*

From this follows a result, which, from its character, belongs to the third part of the present paper, and which may be briefly stated as follows:

Of the m^m possible sets of m numbers which may be chosen from the elements $0, 1, 2, \cdots, m-1$ (with repetition), exactly $N(m)$ are complete residue systems modulo m of some polynomial with integral coefficients.

[*] Compare Part II, end.

Example* 1: $m = 7$. $\mu(m) = 7$; $S(7) = \begin{pmatrix} 7 & 0 \\ 1 & 7 \end{pmatrix}$; $C(7) = \left(\begin{array}{|c|} 7 \\ 7 \end{array}\right)$; $N(7) = 7^7$.

The completely reduced polynomials modulo 7 are the polynomials $a_0 + a_1 x + \cdots + a_6 x^6$, where each coefficient independently assumes all values $0, 1, \cdots, 6$.

Example 2: $m = 2 \cdot 3 \cdot 5 \cdot 11$. $\mu(m) = 11$;

$$S(m) = \begin{pmatrix} 11 & 5 & 3 & 2 & 0 \\ 1 & 11 & 11 \cdot 5 & 11 \cdot 5 \cdot 3 & 11 \cdot 5 \cdot 3 \cdot 2 \end{pmatrix};$$

$$C(m) = \left(\begin{array}{c|c|c|c} 2 & 1 & 2 & 6 \\ 11 \cdot 5 \cdot 3 \cdot 2 & 11 \cdot 5 \cdot 3 & 11 \cdot 5 & 11 \end{array}\right);$$

$$N(m) = 11^6 \cdot (11 \cdot 5)^2 \cdot (11 \cdot 5 \cdot 3)^1 \cdot (11 \cdot 5 \cdot 3 \cdot 2)^2 = 2^2 \cdot 3^3 \cdot 5^5 \cdot 11^{11}.$$

The completely reduced polynomials modulo $2 \cdot 3 \cdot 5 \cdot 11$ are the following: $a_0 + a_1 x + a_2 x^2 + \cdots + a_{10} x^{10}$, where $0 \leqq a_0, a_1 < 11 \cdot 5 \cdot 3 \cdot 2$; $0 \leqq a_2 < 11 \cdot 5 \cdot 3$; $0 \leqq a_3, a_4 < 11 \cdot 5$; $0 \leqq a_5, a_6, a_7, a_8, a_9, a_{10} < 11$.

Example 3: $m = 3^{11}$. $\mu(3^{11}) = 27$,

$$S(3^{11}) = \begin{pmatrix} 27 & 24 & 21 & 18 & 15 & 12 & 9 & 6 & 3 & 0 \\ 3^0 & 3^1 & 3^2 & 3^3 & 3^5 & 3^6 & 3^7 & 3^9 & 3^{10} & 3^{11} \end{pmatrix},$$

$$C(3^{11}) = \left(\begin{array}{c|c|c|c|c|c|c|c|c} 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 3^{11} & 3^{10} & 3^9 & 3^7 & 3^6 & 3^5 & 3^3 & 3^2 & 3^1 \end{array}\right),$$

$$N(3^{11}) = 3^{3(11+10+9+7+6+5+3+2+1)} = 3^{3 \cdot 54}.$$

The completely reduced polynomials modulo 3^{11} are $a_0 + a_1 x + a_2 x^2 + \cdots + a_{26} x^{26}$, where $0 \leqq a_{0,1,2} < 3^{11}$; $0 \leqq a_{3,4,5} < 3^{10}$; $0 \leqq a_{6,7,8} < 3^9$; $0 \leqq a_{9,10,11} < 3^7$; $0 \leqq a_{12,13,14} < 3^6$; $0 \leqq a_{15,16,17} < 3^5$; $0 \leqq a_{18,19,20} < 3^3$; $0 \leqq a_{21,22,23} < 3^2$; $0 \leqq a_{24,25,26} < 3$.

Example 4: $m = 2^7 \cdot 3^4 \cdot 7^2$. $\mu(2^7 \cdot 3^4 \cdot 7^2) = 14$,

$$S(m) = \begin{pmatrix} 14 & 9 & 8 & 7 & 6 & 4 & 3 & 2 & 0 \\ 1 & 7 & 3^2 \cdot 7 & 2^3 \cdot 3^2 \cdot 7 & 2^3 \cdot 3^2 \cdot 7^2 & 2^4 \cdot 3^3 \cdot 7^2 & 2^5 \cdot 3^3 \cdot 7^2 & 2^6 \cdot 3^4 \cdot 7^2 & 2^7 \cdot 3^4 \cdot 7^2 \end{pmatrix};$$

$$C(m) = \left(\begin{array}{c|c|c|c|c|c|c|c} 2 & 1 & 1 & 2 & 1 & 1 & 1 & 5 \\ 2^7 \cdot 3^4 \cdot 7^2 & 2^6 \cdot 3^4 \cdot 7^2 & 2^5 \cdot 3^3 \cdot 7^2 & 2^4 \cdot 3^3 \cdot 7^2 & 2^3 \cdot 3^2 \cdot 7^2 & 2^3 \cdot 3^2 \cdot 7 & 3^2 \cdot 7 & 7 \end{array}\right);$$

$N(m)$† $= 2^{40} \cdot 3^{27} \cdot 7^{21}$.

The following property, which we illustrate only by considering the special case $m = 2^7 \cdot 3^4 \cdot 7^2$, also holds for all m, as is immediately seen. For complicated moduli, it may be used as a convenient check.—On the same linear scale measure off, starting always from the same point, 0, the lengths 2, 4, 6, 8 $= \mu(2^7)$; 3, 6, 9 $= \mu(3^4)$; 7, 14 $= \mu(7^2)$. We obtain in this manner the following marks on our scale: 0 2 3 4 6 7 8 9 14, that is, the first line of $S(2^7 \cdot 3^4 \cdot 7^2)$, in reversed order; and the set of intervals between successive points, 2, 1, 1, 2, 1, 1, 1, 5 gives the first line of the characteristic, in proper order. The analogy with the method of the "Sieve of Eratosthenes" is obvious.

* Compare examples of § 4.

† The fact that each exponent is divisible by the corresponding base, is the expression of a simple theorem which holds for all values of m.

The completely reduced polynomials modulo $2^7 \cdot 3^4 \cdot 7^2$ are $a_0 + a_1 x + \cdots + a_{13} x^{13}$, where $0 \leqq a_0, a_1 < 2^7 \cdot 3^4 \cdot 7^2$; $\cdots$; $0 \leqq a_8 < 3^2 \cdot 7$; $0 \leqq a_{9,10,11,12,13} < 7$.

II. Residual congruences and Kronecker modular systems

§ 7. Residual congruences and Kronecker* modular systems

From Part I, §§ 2, 3, 4, it follows that the existence of a residual congruence $f(x) \equiv \phi(x) \pmod{m}$ is equivalent to the existence of an ordinary identity† in x of the following kind, in which $c_0(x)$, $c_1(x)$, $\cdots$ denote polynomials in x with integral coefficients (which may reduce to constants, including 0):

$$\begin{aligned} f(x) = \phi(x) + c_0(x) \cdot g_0(x) + c_1(x) \cdot g_1(x) \\ + \cdots + c_{\tau-1}(x) \cdot g_{\tau-1}(x), \end{aligned} \tag{6}$$

where

$$g_k(x) = \frac{m}{d_k} \cdot \prod_{i=0}^{\mu(d_k)-1} (x - i) \qquad \text{for } k > 0; \qquad g_0(x) = m.$$

This identity is, in Kronecker's notation, equivalent to

$$\begin{aligned} f(x) \equiv \phi(x) \text{ modd}\Bigg[m;\ \frac{m}{d_1} \prod_{i=0}^{\mu(d_1)-1} (x - i);\ \frac{m}{d_2} \prod_{i=0}^{\mu(d_2)-1} (x - i); \\ \cdots;\ \frac{m}{d_{\tau-1}} \prod_{i=0}^{\mu(d_{\tau-1})-1} (x - i) \Bigg], \end{aligned} \tag{7}$$

and, if we write for shortness $[m;\ \cdots;\ (m/d_{\tau-1}) \cdot \prod_{i=0}^{\mu(d_{\tau-1})-1} (x - i)] = M$, then M is a Kronecker modular system or a Kronecker modulus.

The congruence has this peculiarity, that all functions $g_k(x)$ of the modulus are themselves residually congruent to zero modulo m. The fact that the functions of the modulus determine the complete chain of congruences modulo m, as explained in Part I, corresponds to the combined statements:‡

(*a*) if $f(x)$, $\phi(x)$ are any two polynomials (with integral coefficients) such that $f(x) \equiv \phi(x) \bmod m$, then $c_0(x), c_1(x), \cdots, c_{\tau-1}(x)$ can be determined so that (6) is satisfied;

(*b*) if any function $g_k(x)$ is omitted from the modulus M, then there exist residual congruences $f(x) \equiv \phi(x) \pmod{m}$ for which the identity (6) cannot be satisfied.

In other words: every polynomial with integral coefficients, and whose value is divisible by m for all integral values of x, is representable in the

[*] Instead of using, as in the text, *one* Kronecker modulus, I had originally employed a *set* of congruences with double moduli $(c, \phi(x))$. L. E. Dickson, in conversation, suggested the use of a single Kronecker modulus.

[†] The sign of equality, =, is used, to avoid confusion with the sign for "congruent to."

[‡] On account of the properties (*a*), (*b*), the modular system M may be called a "reduced fundamental" modular system. For a related "fundamental system," introduced for a different purpose, and consequently not reduced, see Hensel, *Ueber die Zahlenteiler ganzzahliger Funktionen*, Journal für Mathematik, vol. 116 (1896), pp. 350–356.

form $\sum_{k=0}^{r-1} c_k \cdot g_k(x)$, and if any of the $g_k(x)$ are omitted, there will be such polynomials $\phi(x)$ which can no longer be represented in the form (6).

Many authors have continued along various lines the work on modular systems inaugurated by Kronecker. One problem, in particular, is the examination of the "equivalence" of two such modular systems, and the reduction of a modular system to a canonical form. This subject is treated, for example, in several papers by Hensel* and Landsberg.* The necessary and sufficient conditions that a modular system shall be in canonical form may, according to these authors, be stated as follows (see 5 of last footnote, p. 365):

A modular system, M, is in canonical form when it is of the form

$$M = [g_1,\ k_1 \cdot g_2,\ k_2 \cdot g_3,\ \cdots,\ k_{r-1} \cdot g_r,\ k_r],$$

where

I. g_i $(i = 1, \cdots, r)$ are polynomials (in x) with integral coefficients and with the coefficient of the highest power of x each equal to unity;

II. The degrees γ_i of g_i $(i = 1, \cdots, r)$ are decreasing integers, $\gamma_i > \gamma_{i+1}$;

III. k_i $(i = 1, \cdots, r)$ are integers and each k_i is a proper factor of the next k_{i+1} (and therefore of all succeeding k);

IV. For $\sigma = 1, 2, \cdots, r-1$, $r > 1$, the polynomial g_σ is divisible by the modular system

$$\left[g_{\sigma+1},\ \frac{k_{\sigma+1}}{k_\sigma} \cdot g_{\sigma+2},\ \cdots,\ \frac{k_{r-1}}{k_\sigma} \cdot g_r,\ \frac{k_r}{k_\sigma}\right],$$

(and this new system is then, as a consequence of I–IV, again in canonical form).

After these preliminary remarks, we may state the relation between Part I of the present work and the theory of Kronecker modular systems as follows:

THEOREM VII: *For a given m the Kronecker modular system*

$$M = \left[1 \cdot \prod^{\mu(m)-1} (x-i);\ \frac{m}{d_{r-1}} \cdot \prod^{\mu(d_{r-1})-1} (x-i);\ \cdots;\ \frac{m}{d_2} \cdot \prod^{\mu(d_2)-1} (x-i);\right.$$
$$\left.\frac{m}{d_1} \cdot \prod^{\mu(d_1)-1} (x-i);\ m\right],$$

*1, 2. Hensel: *Zurückführung der Divisorensysteme auf eine reducierte Form*, I, Journal für Mathematik, vol. 118 (1897), pp. 234–251; II, vol. 119 (1898), pp. 114–130;

3. Hensel: *Ueber die elementaren arithmetischen Eigenschaften der reinen Modulsysteme*, ibid., vol. 119 (1898), pp. 175–185;

4. Landsberg, *Ueber Modulsysteme zweiter Stufe und Zahlenringe*, Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, (1897), pp. 277–303 (277–286).

5. Encyclopédie des sciences mathématiques, Tome I, vol. 2, Landsberg-Hadamard-Kürschák, *Propriétés générales des corps et des variétés algébriques*, pp. 342–366.

See also Hancock, *Canonical forms for the unique representation of Kronecker's modular systems*, Journal für Mathematik, vol. 119 (1898), pp. 148–170.

which is uniquely determined by the signature

$$S(m) = \begin{bmatrix} \mu(m) & \mu(d_1) & \mu(d_2) & \cdots & \mu(d_{\tau-1}) & \mu(d_\tau) = 0 \\ 1 & \dfrac{m}{d_1} & \dfrac{m}{d_2} & \cdots & \dfrac{m}{d_{\tau-1}} & \dfrac{m}{1} = m \end{bmatrix},$$

is in canonical form.

Proof: Conditions I, II, III are obviously satisfied. To show that IV is also satisfied, it is only necessary to translate condition IV into language not involving the special notation of modular systems:

IV. For $\sigma = 1, 2, \cdots, r-1, r > 1$, the polynomial g_σ can be expressed (identically in x) in the form

$$g_\sigma = \psi_{\sigma+1} \cdot g_{\sigma+1} + \frac{k_{\sigma+1}}{k_\sigma} \cdot \psi_{\sigma+2} \cdot g_{\sigma+2} + \cdots + \frac{k_{r-1}}{k_\sigma} \cdot \psi_r \cdot g_\nu + \frac{k_r}{k_\sigma} \cdot \psi_{\sigma+1},$$

where $\psi_{\sigma+1}, \cdots, \psi_{r+1}$, and (as we know), $g_{\sigma+1}, \cdots, g_r$ are polynomials in x with integral coefficients, and (as we know), $k_{\sigma+1}/k_\sigma, \cdots, k_{r-1}/k_\sigma, k_r/k_\sigma$ are integers.

Applying this condition to our modular system M, we obtain for IV:

"We have to show that it is possible to determine, for σ equal to any of the values $1, 2, \cdots, r-1$, a set of polynomials in x, $\psi_\sigma, \psi_{\sigma+1}, \cdots, \psi_r$, with integral coefficients, such that the following equation is satisfied identically in x:

$$\prod_{i=0}^{\mu(d_{\tau-\sigma+1})-1} (x-i) = \prod_{i=0}^{\mu(d_{\tau-\sigma})-1} (x-i) \cdot \psi_\sigma + \frac{d_{\tau-\sigma}}{d_{\tau-\sigma-1}} \cdot \prod_{i=0}^{\mu(d_{\tau-\sigma-1})-1} (x-i) \cdot \psi_{\sigma+1}$$
$$+ \frac{d_{\tau-\sigma}}{d_{\tau-\sigma-2}} \cdot \prod_{i=0}^{\mu(d_{\tau-\sigma-2})-1} (x-i) \cdot \psi_{\sigma+2} + \cdots + \frac{d_{\tau-\sigma}}{d_1} \cdot \prod_{i=0}^{\mu(d_1)-1} (x-i) \cdot \psi_{\tau-1} + d_{\tau-\sigma} \cdot \psi_\tau ."$$

Inspection shows that this happens for

$$\psi_\sigma = \prod_{i=\mu(d_{\tau-\sigma})}^{\mu(d_{\tau-\sigma+1})-1} (x-i); \qquad \psi_k = 0, \quad (k = \sigma + 1, \cdots, \tau).$$

Example: $m = 2 \cdot 3 \cdot 5$.

$$S(2 \cdot 3 \cdot 5) = \begin{pmatrix} 5 & 3 & 2 \\ 1 & 5 & 3 \cdot 5 \end{pmatrix};$$

$$M = [1 \cdot \Pi_{i=0}^{i=4}(x-i);\ 5 \cdot \Pi_{i=0}^{i=2}(x-i);\ 3 \cdot 5 \cdot \Pi_{i=0}^{i=1}(x-i);\ 2 \cdot 3 \cdot 5].$$

Then

$$\Pi_{i=0}^{i=4}(x-i) = \frac{5}{5} \cdot \Pi_{i=0}^{i=2}(x-i) \cdot \psi_1 + \frac{3 \cdot 5}{5} \cdot \Pi_{i=0}^{i=1}(x-i) \cdot \psi_2 + \frac{2 \cdot 3 \cdot 5}{5} \cdot \psi_3$$

for $\psi_1 = (x-3)(x-4)$, $\psi_2 = 0$, $\psi_3 = 0$; similarly

$$\Pi_{i=0}^{i=2}(x-i) = \frac{3 \cdot 5}{3 \cdot 5} \cdot \Pi_{i=0}^{i=1}(x-i) \cdot \psi_4 + \frac{2 \cdot 3 \cdot 5}{3 \cdot 5} \cdot \psi_5$$

for $\psi_4 = x - 2$, $\psi_5 = 0$; and

$$\Pi_{i=0}^{i=1}(x - i) = \frac{2 \cdot 3 \cdot 5}{2 \cdot 3 \cdot 5}\psi_6,$$

for $\psi_6 = x(x - 1)$.

The problem of determining $N(m)$ for a given modulus (see § 6) has its exact counterpart in the case of a general modular system in canonical form, and is treated by Hensel, loc. cit. 3, and Landsberg, loc. cit. 4. In Landsberg's terminology, $N(m)$ is the Norm of the modular system.

(To be continued.)